

29 April 1982

What place for social sciences?

Lord Rothschild's enquiry into the British Social Science Research Council seems to be complete but the report has not yet been written. Here is what he should conclude:

In many places social science is regarded as integral with science itself, under some such name as *Wissenschaft*. In Britain, by contrast, the social sciences have been regarded as studies apart — and as studies that are inherently poor relations of the natural sciences. There has been a research council to support academic research in the social sciences for only the past two decades. During that period the Social Science Research Council has earned a mixed reputation. Like its counterparts elsewhere, it has found that many of the projects it has been supporting can easily be made fun of. It has also — and openly — proclaimed that part of its task has been to improve the quality of academic research in the social sciences. (In the process, inevitably, some of its functionaries have been persuaded to laughable pedantry.) Then, with the expectation that research should be useful, the research council had been asked insistently the old utilitarian question, “what use is it?”. Nobody knows just which of these views persuaded Sir Keith Joseph, the Secretary of State for Education and Science, to let it be known that he is sceptical of the council and most of its works and, soon after his own appointment last September, to saddle Lord Rothschild with the task of pronouncing on the council's future. Perhaps he was simply counting on the likelihood that the man who caused ructions among natural scientists with his recommendations in 1971 for the reorganization of the other research councils would be equally subversive in the social sciences.

What, however, should he say? The first need is to recognize that the abolition of the Social Science Research Council would be intolerable. Whatever its hesitancy in the past two decades and its ways of working, the council has not merely established a number of important and productive academic groups in British universities, but it has substantially achieved that part of its programme intended to improve the quality of research in the social sciences. Moreover, especially at a time when the universities as such are less able than ever before to support research, the abolition of such a serious fund as the research council has become would create much more damage than the universities could sustain.

So, should the Social Science Research Council be reorganized in some way, and if so, how? In the past three years the council has been doing everything it could to make its procedures more understandable to those whose work it supports and at the same time to make its intentions more widely appreciated by the government and in the country at large. In the process the council has angered many of those who had grown used to the old procedures. Thus the council now finds that its critics are in two camps — those who preferred the past, and those who are unconvinced that its present procedures will deliver what people expect of the social sciences.

Nobody should be dismayed that this turn of events has come about. Who asks that a newly created organization such as the council should be able, within a couple of decades, to find a sure

way of supporting research in the social sciences? The problems, acknowledged by everyone concerned at the outset, are too serious for that. And in any case, the successes in the council's portfolio of previously supported projects are quite sufficient to outweigh the hasty criticisms of those who point to projects that have failed. If some reorganization of the social sciences is needed, and it may be, attention should be directed to the present mismatch between what the council does in the support of academic research and the uses which government departments make of academic skill and research in the social sciences.

One of the most valuable parts of Lord Rothschild's declaration on the Social Science Research Council will undoubtedly be the calculation he will have made of the amount of research in universities commissioned directly by government departments and that supported by the research council. Paradoxically, the circumstances are the inverse of those that stimulated Lord Rothschild's recommendation of 1971 for the natural sciences. There, a decade ago, research councils such as those responsible for agriculture and medicine were free to decide for themselves what research strategy should be pursued, while the ministries most directly concerned were able only in passing to offer advice. The recommendation, immediately accepted by the Heath government, that in future (as now) the ministries should commission research in important areas from the research councils was sensible. The fact that the “customer-contractor” principle has not worked entirely satisfactorily is no fault of the principle, but of those who have tried to apply it. In the social sciences, on the other hand, the volume of research commissioned directly by government departments is far greater than that which the Social Science Research Council is able to support. Moreover, on the face of things, the projects commissioned directly by government ministries are put out to academic research groups with comparatively little advice from the Social Science Research Council, the body best equipped to tender an informed opinion. So is there not in the future organization of the Social Science Research Council, a need that the whole volume of public support for research in the social sciences should be channelled through some common agency?

The question remains of what should be the relationship between the social and the natural sciences. Academics, notoriously hide-bound, have over decades argued the toss about the place that social science should occupy in academic life. Economics is accepted, now, as an inevitable part of the spectrum of what goes on in universities. More esoteric subjects — anthropology on the one hand, sociology on the other hand — are more easily scorned, but wrongly. Physicists may hold that the data on which people in these disciplines base their arguments, and the arguments themselves, are fuzzy by the standards that apply in, say, the analysis of data collected in some objective measurement of the real world. But was there not a time when even physics was a fuzzy field of speculation? Can it be right that the natural sciences should continue to laugh at the difficulties with which the social sciences must contend? For what the field needs is what the Social Science Research Council has been saying for the past twenty years is an essential part of its own brief — the general improvement of methods of teaching as well as of research that will, in due course, make social science accessible to a wider circle, and thus more useful.

US contributors please note

From now on, authors in the United States and Canada should send their manuscripts either direct to London or to the Washington office, whose new address in *Nature*, 991 National Press Building, Washington, DC 20045; manuscripts should no longer be sent to *Nature*'s New York office.



'Freeze' frenzy hits US

Does the public demand for a nuclear weapons 'freeze' reflect a basic shift of opinion?

Enthusiasm for a freeze on the production of nuclear weapons is sweeping the United States. Rallies and teach-ins abound, with participants waving pictures of Hiroshima victims and charts of death from "limited" bomb blasts. Students and faculty on many campuses are involved. So are physicians. Even in the southern states, which was fairly unmoved by the Vietnam protests, freeze frenzy is taking hold.

The movement has been accompanied by pronouncements from prominent people calling on the Reagan Administration to stop its weapons build-up, and blaming the arms control and national security community for giving the public gobbledygook and no results for its 20 years of effort. Historian Barbara Tuchman, whose ability to understand the intricacies of history has won her two Pulitzer prizes, threw up her hands in pages of the *New York Times* recently and confessed that "the subject of nuclear arms control . . . is virtually incomprehensible to the layman". She showed how arms control has failed since the beginning of the century. Meanwhile, Yale psychiatrist Robert Jay Lifton has blamed the intellectual community for not doing enough. "This central issue of our times has been fundamentally ignored in the universities", he said at a recent conference jointly sponsored by two education groups and Hobart and William Smith College. The failure of leaders and the security community was bemusedly outlined in public recently by a former participant, Roger Molander who after years working on nuclear arms control in the White House quoted a presidential science adviser as saying, of the leaders' behaviour "where are the grownups?" Molander quit and founded "Ground Zero", a grass-roots organization that has been remarkably successful in sparking rallies around the land.

Are we witnessing a fundamental shift in US opinion? Europe is watching it closely, particularly the European peace movement, which is hoping that it will dissuade the Administration from its hawkishness, and perhaps tip the balance against NATO plans to install new missiles in Europe.

But the shift may not be fundamental. It is caused by two things, both of which can change quite soon. First, the President and Mr Caspar Weinberger, Secretary of Defense, and other officials, have been saying things in public that previous Administrations were wise enough to avoid. They have talked openly, and casually about the likelihood of nuclear war, the "fact" of Soviet military superiority, the need for a gargantuan, apparently ceaseless weapons build-up, and they have appeared patently insincere about arms control. By escalating public discourse, clearly they are upsetting people.

The second cause of the freeze frenzy is the volatility of public opinion itself, which is now swinging back from its hawkish mood of recent years. The last time a serious arms control plan was presented to the public, SALT II in 1979 and 1980, they were clearly bored. Then the humiliation in Iran, and Soviet invasion of Afghanistan, sparked a conservative trend which eventually brought Reagan to office. If the public is asking why, today, there is not sufficient arms control, it has itself, in part, to blame.

European peace movements should realize how different the American freeze movement is. The US demonstrators seem not too concerned about the proposed emplacement of Pershing and cruise missiles in Europe — the 1979 NATO decision which Europe's peace movement wants reversed. To the Americans, those Pershing and cruise missiles are small fry compared with Reagan's plans for the MX missile and an apparently ceaseless parade of giant aircraft carriers, submarines, nuclear-armed B-1 and Stealth bombers, and the like which the American movement seeks to freeze. Nor is the neutron bomb a major US issue. The two movements have not fused, although some major gaffe by the Reagan Administration could yet cause it to happen.

And Americans should question whether their intellectuals are

to blame — whether, as Lifton grandly said, "a generation of scholarship has been lost" — whatever that means. To blame the arms control and national security intellectuals for the military build-up in the world is like blaming social scientists for the persistence or racism, poverty, and violence. True, one of the vanities of intellectuals is that they think that by studying a subject, they can control it. But not even scientists can do that. William Perry, former Under-Secretary for Defense Research and Engineering, and one of the key defence intellectuals who tried to "sell" SALT II for the Carter Administration, says there is "some merit" in the view that US defence intellectuals have failed the country. But it was the public's fault too, he notes "very little else at the time (besides SALT II) was guaranteed to put the public to sleep". Perry welcomes the freeze movement — although he disagrees with some of its aims. "As a leader in an Administration I would rather work with an aroused public, and try to channel that interest, than have to try to convince an apathetic public."

It would be unwise for the Reagan Administration to view the freeze frenzy as a suspect foreign import, festering on US university campuses, and accuse its adherents of lack of patriotism. While the movement has set itself against the Administration's arms build-up, it would clearly be delighted by any presidential move to talk seriously about arms control with the Soviet Union. For the first time in years arms control is gaining a public constituency in the United States. It is too bad that SALT II is not before the Senate, so that the United States, Europe, and the Soviet Union, could benefit from it.

Hasty abolition

Disappearance of the Schools Council was predicted, but may have been unwise.

The British Secretary of State for Education and Science, Sir Keith Joseph, is plainly an abolitionist by temperament (see preceding page). Hankerings to get rid of the Social Science Research Council are one thing. Last week, however, he abolished the Schools Council, an organization set up in Britain in the 1960s in the hope of overcoming some of the obstacles to national planning in British schools education. For in Britain, as elsewhere, the central government has no direct authority over teachers in local schools, who are employed by and responsible to their local education authorities. The result is that the curriculum can vary enormously from one place to another; school leaving examinations can differ in stringency from one examinations board to another; and in principle there is no reason why some local authorities should not give their schools entirely eccentric marching orders. The problem of bridging this gap has daunted many governments and yet is politically insoluble — the local authorities in Britain would not sacrifice such independence that they retain in this small area of their operations.

The Schools Council was designed principally as a talking shop, a place where central government, local government and independent people could be represented and in which, tacitly at least, local authorities and the central government would come to some understanding on education policy. What has gone wrong is that the Schools Council's procedures have become cumbersome and the council itself has been politicized by the conflicting interests it represents, often those of a teachers' labour union. At the same time, the original plan to bring about some coherence in British schools education by advising novel and more stimulating curricula has been defeated by the shortage of funds and, to some extent, by the apathy of the schools. The Schools Council has in the past few years been a sitting duck, ripe for abolition. Sir Keith Joseph has done this deed and now plans to replace the Schools Council by two advisory bodies, one for the curriculum and one for examinations. The snag is that these advisory bodies will not, by their constitution, bring central and local government together. When he is more experienced in his present post, Sir Keith Joseph will regret this.

US biomedical research threatened

New director at NIH faces discontent

Washington

The primary sponsor of US biomedical research, the National Institutes of Health (NIH), is at last, after a ten-month hiatus, getting a new director. He is Dr James B. Wyngaarden, whom the Administration and the biomedical community agree is a good choice. Wyngaarden had his Senate confirmation hearing last week and is expected to be confirmed and sworn in within the next few weeks. But he will be inheriting a storm.

When Wyngaarden moves in to NIH's blossoming campus on the fringes of Washington, he will have to face the anger of deans and administrative officers of the 2,500 medical schools and universities in the United States who receive NIH funds and who are furious with NIH for proposing to cut their indirect cost allowance by 10 per cent in 1983. Although the immediate sum involved is a trivial \$70 million in NIH's proposed \$3,700 million budget, the administrators fear that the cut portends a fundamental shift in the structure of the US university economy. Whether it does or not depends on Congress, which may yet restore the money.

Wyngaarden, 57 and a Democrat, is Hanes Professor of Medicine and chairman of the department of medicine at Duke University School of Medicine in North Carolina. Among other Washington assignments, he has served on the President's Science Advisory Committee. His nomination was backed by NIH's many supporters on Capitol Hill, and by Richard S. Schweiker, the chief of NIH's parent Department of Health and Human Services, who is himself a former congressional backer of NIH. Although some of the Reagan Administration's key science appointees have included unknowns or outsiders to the basic science community, Wyngaarden is neither.

Wyngaarden's first job will be to find directors for the five NIH institutes which are without permanent directors, some since January 1981. Another key post, that of deputy director of NIH for science, is also vacant. The previous director, Donald Frederickson, drew up lists of candidates last July, and Wyngaarden will have to see how many candidates are still interested in the jobs. In federal research agencies, when a permanent chief is lacking, normal business continues. What suffers is the organization's internal needs and long-term planning.

Wyngaarden inherits a dispute over indirect cost allowances that is dividing university administrators and faculty across the country. The bulk of NIH funds for research project grants and research centres is awarded in two parts. The first part is the direct cost of the research — salaries, equipment, travel, publications, and the like. The second part covers indirect costs, a percentage of the direct costs that the government agrees to pay to the researcher's home institution to keep him housed. Indirect costs include such things as central fuel and electricity bills, cleaning, library and central animal facilities.

Previously NIH reimbursed the institutions a fixed 15 per cent of a researcher's direct costs, no matter what they were. However, this was deemed arbitrary and NIH allowed the percentage it would repay to grow, frequently reassessing it to take account of rising utility costs and inflation. For 1983, NIH estimates that indirect cost payments on research grants and centres will average 30 per cent of the direct costs. Other federal agencies that fund research follow the same process but with percentages that vary among agencies and universities. (The

resulting patchwork of repayments would be unbearable for university administrators were the funds not essential for their hard-pressed institutions.)

Thus, while the growth in research budgets has tended to slow in recent years, indirect costs have kept on rising. A Columbia professor estimated recently that fifteen years ago his research supported by NIH enabled Columbia to collect \$2,000 in indirect cost repayments. During that time, even though his research effort has only doubled, Columbia has been able to collect \$40,000 in indirect cost reimbursements on it.

However, the array of administrators and their representative societies, such as the Association of American Medical Colleges and the Association of American Universities who oppose the NIH plans to reduce its contribution to administrative costs, are not merely concerned about paying next year's heating bills. They claim that NIH is shifting the whole debate about what the institutes are for and how they should maintain US world leadership in biomedical research.

During his six years as NIH director, Frederickson argued that researchers suffered from uncertainty as to what would

Franco-Soviet space flight plea

Three pressure groups of French scientists and intellectuals, "Comité des Physiciens Français", "Comité Sakharov" and "Solidarité au Spacenaute", have called for France to withdraw from the Franco-Soviet manned spaceflight planned for June.

Following the procedure of the manned Interkosmos flights, which included representatives of all Comecon block countries, two French candidates for space — Jean-Loup Chretien and Patrick Baudry, are now training in the Soviet Union. One of these will be selected to be launched with a Soviet partner to rendezvous with and work aboard the Salyut-7 space station put into orbit last week.

The pressure groups base their demand on the recent increase of human rights violations in the Soviet bloc — the invasion of Afghanistan, the banishment of Sakharov, martial law in Poland. For a relatively minor scientific gain, they say, France would be reduced to a "supernumerary" in a Soviet public-relations spectacular. The propaganda value to the Soviet Union of a joint flight involving, for the first time, a guest from outside the Socialist bloc, would be immense.

For all three groups agree that at this stage France stands to gain little from the proposed manned flight. Franco-Soviet cooperation in space goes back to the visit of President de Gaulle to Baikonur in 1966,

and over the years has included a laser ranging experiment on the lunokhod Moon-rover, gamma ray spectrometers carried aboard Venus probes, and the launching by the Soviet Union of the French Aureole geophysical satellites. French participation is also planned in the Soviet mission to Venus and Halley's comet in 1985.



The proposed French participation in the Salyut mission would cover three main fields, astrophysics, materials sciences and biology. Comité de Physiciens Français, in particular, contends that the first two programmes could be carried out equally well aboard an unmanned automated spacecraft. As for the biology programme, this will simply allow the "well-known" physiological stresses of space to be applied for the first time to a French citizen.

Vera Rich

happen to the research budget from one year to the next. This discouraged bold thinking and long-range research plans, he said.

According to Frederickson, NIH should fund 5,000 new "noncompeting" research project grants, in addition to keeping previous years' commitments to projects lasting two and three years. In this way, researchers would know there would be 5,000 research project grants available for the best proposals. As a corollary, Frederickson proposed that training grants should be stabilized, with the budget allowing for 10,000 of these, each year. The numbers were arbitrary and widely known to be a budget ploy, but they had the salutary effect of keeping everyone's attention on a few key concepts for the health of research: stability, availability of money for new projects, encouragement to young researchers.

Sources close to the NIH budget process say, however, that last December, when NIH received a 1983 budget figure, it calculated that it would allow funding for only 3,500 new research project grants, embarrassingly far below the 5,000 goal publicized by Frederickson, and well below the 4,741 new project grants funded in 1982. Schweiker therefore told NIH to change the budget, keeping the same total, so that the number of new project grants would be above 4,000. As a result, NIH proposed taking the cut in the administrative side, in direct costs, to avoid an enormous fight with scientists claiming that NIH was abandoning its commitment to innovative research.

So now the universities and research institutions are asking why they should be singled out for cuts — and what Pandora's box has been opened for future years.

According to John F. Sherman, former deputy director of NIH and now vice-president of the Association of American Medical Colleges, which protested to Schweiker in March, it is likely that the 10 per cent cut proposed for 1983 will set a precedent for unpredictable cuts in NIH's indirect cost allowances in future years, and perhaps lead to a lowering of indirect cost allowances by other federal agencies.

"If there is anything to divide faculty from academic administrators, this is it", says Sherman. "The institutions' representatives are frantic. They see no solution except for Congress to add the money. Whereas the faculty are more willing to see indirect cost reimbursements cut in order to maintain a higher number of competing research awards."

If NIH's indirect cost reimbursements are cut, however, institutions may be forced to reduce the size of their research programmes. And if NIH awards fewer research project recipient grants, the institutions will have a smaller base for seeking indirect cost payments. To hear them talk now though, the two sides seem not yet to have realized that they have a common cause. **Deborah Shapley**

Commercialization of research

Patent views

New York

Close on the heels of the closed meeting of five prominent university presidents at Pajaro Dunes, California, last month (*Nature* 8 April, p.479), the first of many other, open meetings was held last week in New York to discuss the growing issue of commercialism in academic research. This one was sponsored by the New York Bar Association's patent committee. Participants included Dr Steven Muller, president of Johns Hopkins University, Joshua Lederberg, president of Rockefeller University and A. Thomas Bartlett, president of the American Association of Universities (AAU).

The participants concurred with the conclusions of the Pajaro Dunes meeting concerning the problems of the commercial use of university-based genetic engineering research. In the words of Dr Muller, "the hazards to universities are manageable". It seemed clear at the meeting, though, that there can be no sweeping, general rules covering every campus and field of research in the nation, and that university policies will be built up step-by-step.

Those at the bar association meeting felt that university faculty would be well able to defer publication of key research in the open literature to satisfy patent needs of the commercial partners. Dr Lederberg noted that while such delays are "natural hindrances" to normal university activity, they should be acceptable to faculty so long as delays are reasonable. Dr Muller pointed out that Johns Hopkins recently passed regulations requiring university researchers to publish in the open literature and to discourage undue secrecy while faculty file for patents. The new regulations also place the burden of proof on faculty members who claim that a private invention has no relation to their university work.

One problem highlighted in the discussion was that European visitors to US campuses could pick up innovations in the course of seminars and talks with US scientists, go back to Europe and patent them, leaving the US academic inventor of the idea with no protection. This possibility arises because European law awards patents to the first to file for one, who may not necessarily be the inventor as in American law.

There was a call at the New York meeting for a re-examination of the relationship of professor to university and the traditional rights of university employers and their faculty employees, often ignored in faculty contracts. For example, can a university pressurize one of its professors to pursue a patent if he is reluctant to do so?

Dr Lederberg encouraged universities to help their faculty negotiate their outside contracts and to define extramural and intramural roles. He was concerned that

many faculty are under-selling their services, and recommended that they charge six times their academic rate of pay for outside consultation. The universities can benefit, he said, by arranging that these fees be shared equally between a faculty and member and the institution.

Finally, several details of the best known university-industry agreements are now available to university presidents seeking to negotiate their own industry arrangements. University presidents seem to be making a practice of circulating current or draft agreements with industry among themselves for purposes of comparison.

Last year, Representative Albert J. Gore (Democrat, Tennessee) attacked the universities for their new industry ties and asked AAU to establish general guidelines on the matter. But the New York meeting seemed to confirm that such general guidelines are unlikely to emerge, or be seen as practical. As AAU president Bartlett said, "we are groping our way into the future". **Michael D. Stein**

Dutch pharmaceuticals

New era vaccine

Waalre, The Netherlands

Intervet International, a Dutch veterinary pharmaceutical company, claims to have marketed the first vaccines produced through recombinant DNA technology. The Dutch are proud to have beaten two American companies — Norden Laboratories and Cetus — which recently announced that they were developing a vaccine against pig scours for introduction on to the market within two or three years. Intervet marketed two scour vaccines at the end of March.

The vaccines are intended to prevent infectious diarrhoea of pigs and calves, a major cause of which is the *Escherichia coli* bacterium. The bacterial components implicated in causing scour in pigs and calves are the adhesion factors K88 and K99 respectively. Combination of these factors with a special adjuvant was used by Intervet to produce a vaccine which, when given to the sow or cow, causes the animal to produce protective antibodies which are passed on to their offspring through the colostrum or first milk. In the past adhesion factors K88 and K99 have been purified from wild strains of bacteria isolated from the field. Using cloning techniques Intervet transferred the genes responsible for the production of the K88 and K99 into a laboratory strain of *E. coli* (K-12) which then produces a much greater quantity of antigen for the preparation of the vaccines.

The work to produce these strains started in 1979 in collaboration with the Dutch Institute of Health and was later transferred to the DNA research group within Intervet's mother-company Akzo Pharma, itself part of the Dutch chemical multinational Akzo. **Casper Schuurings**

Environmental research funding

Bleak outlook

The UK Natural Environment Research Council (NERC) is struggling against the somewhat cool attitude of the Minister of State for the Environment, Mr Michael Heseltine, towards science. The Chief Scientist's office at the Department of the Environment (DoE) is doing its best to protect NERC from budget cuts, but the task is not proving easy.

In fact, the department's attitude was so discouraging late last year that even NERC's chairman, the sanguine Sir Hermann Bondi, was beginning to be worried. When Bondi took up his post in October 1980 he defended those of Heseltine's cuts which were already affecting NERC, in the name of the "Rothschild principle" — that the customer (in this case DoE) pays for what it wants and the contractor (NERC) provides it. After more than a year's experience of DoE attitudes, however, Sir Hermann's line has softened a little: "I still believe in the Rothschild principle" he says, "but Rothschild assumed that there would be enlightened customers taking a long view".

The problem for NERC does not seem to lie in fundamental, or in applied research, but in the "strategic research" which is not obviously one or the other. A prime example is the geological survey of Great Britain, whose science may appear to be uninspiring, and which is applied only occasionally for major construction projects. In the long run, however, it provides a basic data set without which much other work would not be possible.

DoE's attitude has pushed NERC into finding more commercial contracts both in the United Kingdom, and abroad; but, say NERC staff, contracts in the United Kingdom cannot supplant the geological survey because they are too patchy, too specific and often entail a degree of commercial secrecy. Moreover, British firms are at least as short of cash as the government at present. Abroad, NERC have been seeking contracts actively, but without any major successes. Bondi would like to encourage this kind of expansion, but NERC is relatively new in the field and must compete with seasoned operators like the US Geological Survey and the Institut du Physique du Globe of Paris.

In the meantime, NERC staff must be supported — and the problem has been compounded by DoE delays in defining a programme of research for 1982–83. By early April, NERC would normally have expected to know DoE's requirements for the year. Worse, the council has been told not to spend money on the basis of DoE "letters of intent" until everything is signed and sealed, a further break with the usual practice.

However, there is some light at the end of the environmental tunnel. Recently NERC held a top level meeting with its three main

customer departments, those of energy, industry and environment and with the Central Policy Review Staff (the Cabinet's think tank) and there was a consensus that the problems lay with strategic research and long range planning.

NERC staff argued that if only they could be told the requirements of the departments well in advance, then the council could design a good basic and strategic research programme which would also yield the contractual results that the departments wanted. So the departments will work together to produce a five-year forward look for NERC, with the intention that this would be "rolled forward" annually thereafter.

Robert Walgate

NERC stays afloat

The British government has allowed the Natural Environment Research Council (NERC) one present: a new ship, specially designed for research, to replace *Shackleton* which is now a creaking 28 years old. The vessel is being bought on a loan from a merchant bank, and will cost NERC about £1 million a year for eight years out of its current £56 million budget.

The ship will be called *Charles Darwin*. It will not be ice strengthened, but will have a clear aft deck for equipment and will be able to tow over the stern, a big simplification over *Shackleton* (a converted container ship) which has to pull equipment from an A-frame over the starboard side. The *Charles Darwin* will be fitted for general research from marine biology to geophysics, and will provide the largest on-board laboratories of the NERC fleet. It should be delivered in 1984.

The research fleet will then include: *Discovery*, also a general purpose research ship, which must retire around 1989; the marine biological ships *Frederick Russell* and *Challenger*; and two ice-strengthened vessels. There are three other English research ships owned by the Ministry of Agriculture, Fisheries and Food (for fisheries research).

Robert Walgate

US recombinant DNA guidelines

Expected change

Washington

Continuing the trend towards relaxing controls on the newly commercial business of recombinant DNA research, the National Institutes of Health (NIH) has published revised guidelines that govern most US academic research and some industry research in the field. There are few surprises in the new guidelines; they follow the plan voted by NIH's Recombinant Advisory Committee (RAC) meeting in February (see *Nature* 4 February, p.358).

The revised guidelines (published in the

Federal Register, 21 April and effective from that date) contain three important features. First, although there was pressure to abandon the present system of mandatory controls and change to a voluntary system, the mandatory system was retained. This means that organizations performing recombinant DNA research must still each have an institutional biosafety committee, and, for those doing work at P3 and P4 containment levels, a biological safety officer.

Second, no class of experiments is "prohibited" any longer. Of the five classes of experiments that were prohibited under the old guidelines, three are still required to undergo RAC review and NIH approval before initiation. These are drug resistance traits, toxin genes and deliberate release to the environment. Third, the section of the guidelines dealing with containment levels has been drastically simplified. Many experiments previously requiring NIH approval now require only the approval of the local biosafety committee.

In approving these guidelines, Richard M. Krause, Director of the National Institute of Allergy and Infectious Diseases of NIH, rejected a proposal for far greater relaxation of controls put forward last April by Dr Allan Campbell and Dr David Baltimore. These would have required a major overhaul of the guidelines. Apparently, Krause agrees with the majority of the members of RAC that such an overhaul is not appropriate — yet.

Deborah Shapley

UK nuclear power

Cracks no worry

New design criteria for the pressure vessel for a British pressurized water reactor (PWR) has been given the cautious approval of Britain's most distinguished PWR opponent, metallurgist Sir Alan Cottrell, now Master of Jesus College at the University of Cambridge.

The new criteria are the result of the work of the United Kingdom Atomic Energy Authority study group, headed by Sir Walter Marshall, on the integrity of PWR pressure vessels. The group last reported in 1976, but six years later the situation requires a fresh look. The design of the Sizewell PWR is going ahead; there have been changes in methods of steel manufacture; and there is now an accumulation of evidence and experience of stress corrosion cracking and of catastrophic failure tests.

In an open letter to Sir Walter Marshall, Sir Alan writes: "I would expect us to avoid the problems of localized cracking, underneath the stainless steel cladding, which the French have experienced in some of their vessels".

He adds "Whether we shall be entirely clear of the problems of metal fatigue and

crack growth associated with corrosion is less certain, but the evidence now is far better than it was in 1976 and I agree with the report and you that it shows that crack growth rates are — or can be, if the quality of the metal and the water chemistry are closely controlled — much lower than appeared to be the case in 1976”.

The report estimates that a hidden crack in a pressure vessel would have to be at least 75 mm in “height” to initiate crack propagation under normal operating conditions. (“Height” is defined in the report to be the vertical distance from the surface between the nearest and furthest surfaces of the crack.) The best estimates of stress corrosion crack growth indicate that cracks of a starting height of, say, 50 mm, would grow less than 10 mm during service.

The report also describes methods by which cracks can be prevented and monitored during the manufacture of the pressure vessel and during service, and in these respects, says Cottrell, the report goes much farther than its predecessor.

However, in the end Sir Alan remains very cautious. “To summarize”, he writes “provided the recommendations in the report are all accepted and fully applied in practice; and in particular that arrangements are made to keep all parts of the pressure vessel in the ‘upper-shelf’ range of fracture toughness at all times; and arrangements are made to ensure that the vessel is free at all times from cracks, down to a size at which growth becomes insignificant under all conditions; then a PWR pressure vessel subject to all these conditions will have a high integrity and reliability in service”.

Robert Walgate

University redundancies

Now the crunch

British universities seem well on the way to solving their budgetary problems by persuading tenured academics to opt for early retirement. Hundreds of applications to retire early from members of teaching staffs have already been accepted. But many universities are left with the nagging doubt that the terms agreed with academics willing to go early may not qualify for reimbursement under the scheme the University Grants Committee (UGC) has now made public.

In February the UGC confirmed that it would reimburse in full the cost of payments to staff made redundant because of cuts in the government grant. In the UGC's letter to the universities, the term “redundancy” covered early retirement. For academic staff over 50 the approved terms are those of the Premature Retirement Compensation Scheme (PRCS) of the University Superannuation Scheme (USS) introduced in 1979. An eligible employee receives the USS pension and any accrued benefits plus a compensatory pension and lump sum payment calculated in “Scheme” years. The cost to the university

is the cost of paying into the pension fund this extra benefit.

The scheme allows academic staff below the age of 50 — not covered by PRCS — to receive a deferred pension, a lump sum payment of one month's pay for every year of service, and one month's pay for every year of service in excess of five years or after their 30th birthday, whichever is the greatest. Payments of up to £55,000 plus pension rights could be made to a redundant 50 year old. The scheme is essentially that put forward by the Committee of Vice-Chancellors and Principals except that the pensions of staff aged 50–55 who opt for early retirement will not be index-linked.

Conditions for reimbursement are stringent. The PRCS was not formulated as a redundancy compensation scheme and universities offering enhanced terms will not be repaid. Terms negotiated before 5 February will, however, be supported but only up to the limit of the government-approved scheme. Staff cuts must also be “consistent with academic plans” — UGC may refuse to pay out for redundancies which do not form part of a comprehensive economy drive.

Volunteers for redundancy schemes will have a good chance of being re-engaged for part-time teaching. The UGC will not reimburse for staff who are to be replaced, but recognizes the need to make a gradual adjustment to staff losses. The cost of taking back staff on a part-time basis will be reimbursed only if contracts are signed before the end of the academic year 1984–85 and do not extend beyond three years.

How are the universities coping with the cuts? Salford is hoping to lose 500 staff members (120–145 of them academic) by the end of the academic year 1984–85. Twelve of its 45 professors are to leave by September, and in order to maintain a balance of experience and age distribution the university is keen to get voluntary severances from younger staff.

Leeds was one of the first universities to respond to government cuts by planning staff losses and it is hoping that its enhanced retirement scheme will make it eligible for compensation. Of the 145 applicants, 95 will have been accepted for early retirement. The bulk of staff losses were accepted under the first stage of its economy plan — a 6 per cent cut in expenditure by the end of the academic year 1982–83, but cuts to its budget and the repercussions of the government's overseas students policy have now forced a 10 per cent reduction in expenditure.

At the University of Cambridge there are doubts as to whether the university's plans will be eligible for UGC compensation. Forty academics have responded to an appeal to take early retirement. A scheme providing enhanced compensation was offered to university staff aged 58 to 60 after its approval by the Council of the Senate in February. The university hopes to attract 100 volunteers in a drive to knock £2 million a year from its expenditure by

July 1984. The university has had to offer enhanced terms because, compared with the age of 65 on which the PRCS scheme is based. Cambridge academics have tenure until 67. Eligible applicants will receive a full pension plus a lump sum calculated on the basis of the pension they would have received at normal retirement. The university is pleased with the response so far.

Jane Wynn

UK biotechnology

A start in sight

The British government has at last fulfilled its promise, made more than a year ago, to establish machinery for coordinating the national effort in biotechnology. Last week, Kenneth Baker, the minister for information technology, announced the setting up of an inter-departmental committee on biotechnology under the chairmanship of Dr Ron Coleman, the recently-appointed Government Chemist.

Dr Coleman's chairmanship may seem rather unusual, chiefly because the Laboratory of the Government Chemist of which he is head, has traditionally followed a fairly narrow remit of providing analytical services to the public sector. But Dr Coleman's desire to broaden its scope has persuaded the Department of Industry to make the laboratory responsible for the department's interest in biotechnology. The new committee's membership includes all those government departments with an interest in biotechnology together with the medical, agricultural and science and engineering research councils, the British Technology Group, the Public Health Laboratory Service and the Centre for Applied Microbiology and Research, Porton Down.

The chief task of the committee, which will have no money of its own, will be to coordinate its members' separate interests. One question close to its heart will be how to transfer the results of biotechnological research into industry. That will involve paying special attention to the role of the Biotechnology Directorate, recently set up by the Science and Engineering Research Council, and the British Technology Group, whose task it is to exploit academic research in industry.

The committee's other main objectives are to foster international collaboration in biotechnology and to look for ways of ensuring the continued existence of culture collections, which at present suffer from being funded from a variety of different sources. The committee, which has before it a study commissioned by the industry department on the organization and funding of culture collections, will be looking for one government body to take over their administration and may recommend changes to the fees now charge for the services provided by the collections.

Judy Redfearn

French universities

Into reverse

One of the first steps towards undoing the authoritarian reforms of the French university system introduced by the previous minister of universities, Alice Saunier-Séïté, has just been taken. Students have been allowed back onto the university councils, the senior administrative councils of the universities.

The reform would be radical — if the students were radical. But the present atmosphere is nothing like that of 1968. The new student representatives are conservative and unreforming, say university staff. For example, in one recent meeting in the University of Paris, students warmly approved a proposal for more examinations and greater selectivity — something which would have been anathema fourteen years ago.

The underlying question is not what the students will do, but where the government will stop in this new attempt to restructure French higher education. M. Alain Savary, the minister for national education (which now includes the universities and *grandes écoles*), has some plans which are causing shivers in the French establishment. Not least among them is that the *classes préparatoires*, the usually private courses that prepare students for entry into the prestigious *grandes écoles*, should be put into the hands of the universities.

This would mean that the brightest of French schoolchildren, who normally pass through the *classes préparatoires* on their way to the *grandes écoles* (and thence to management and government administration), would be exposed to a university environment. There they would learn that the universities were not the second class establishments that some suppose, the argument goes, and some might even choose to stay and read for university degrees.

However, such proposals are too radical for the *grandes écoles*, who are resisting them so strongly — and are so well represented in government — that the political will at the ministry of education is beginning to falter.

For researchers, the situation is even more complicated. There is a conflict afoot between the minister for research and technology, Jean-Pierre Chevènement, and M. Savary. Chevènement has battled long and hard for control of all government spending on civil research and development. Much of the research is done in the university environment, in laboratories which are now funded by Chevènement's ministry. In general, the researchers in these laboratories (such as those of the Centre Nationale de la Recherche Scientifique) do little formal teaching in the university (although they help to train researchers through seminars and research degrees). Chevènement is believed to want more of the teaching to be done by his researchers,

to improve the training of scientists; but this would remove some of the university teachers' higher level and more interesting work. So university opposition is growing to these moves, and Savary inevitably represents one side and Chevènement the other.

How this battle will develop is anyone's guess, but it could easily lead to another deadlock, when the promised higher education reforms would merely tinker with the system, rather than removing its two main faults: the class divisions between the *grandes écoles* and the universities, and between the researchers with formal teaching responsibilities and those without.

Robert Walgate

Moves on energy

The French government is showing every sign of pain in its attempts to embrace the environmental movement. In the latest agonies, the council of wise men which advises the Minister of Industry on nuclear safety has been reformed to include a few token nuclear opponents; and the promised new agency for renewable energies is to be set up, but shorn of FF 170 million.

The safety council is the Conseil Supérieur de la Sûreté Nucléaire, presided over by Professor Louis Néel, an ex-director of the French atomic energy agency laboratories in Grenoble. Néel is described by his opponents as "totally pro-nuclear" and, indeed, his council has met only three times in the past seven years, giving the previous government a pretty clear run on safety matters. Under the new arrangement, Néel stays as president of the council, but it is to meet more frequently (two or three times a year) — and there will be a few cuckoos in the nest.

These come in part from the addition of a further seven experts in the council, including a science journalist, M. Serge Berg of Agence France Presse. The biggest change, however, is the inclusion of trade union representatives on the council for the first time, among them Bernard Laponche, a prominent nuclear critic from the more environmentally inclined of the two major unions, CFDT. So far Laponche is not impressed. The council is still not complete and no date or programme has been set for its first meeting.

As for the renewable energy agency, Agence Nationale pour la Maîtrise de l'Energie (ANME), this will regroup the existing solar energy and energy conservation agencies, and control their budgets; but present indications are that its budget will be less than the sum of its predecessors. The cut in budget is largely the result of a deal to buy Algerian natural gas — the energy conservation agency loses FF 170 million (£17 million) to help pay for the gas.

Robert Walgate

Radioactive waste in India

Looking to glass

New Delhi

The world's second commercial plant for the immobilization of radioactive wastes from nuclear power generating stations is soon to be built at Tarapur, in the western Indian state of Maharashtra. The first such plant is already in operation at Marcoule in France.

The new plant is being set up in the wake of increasing concern in India that radioactive waste from nuclear power installations might damage the health of the people — although there has been no scientific evidence of such damage so far from India's two nuclear power plants at Tarapur and Kota in Rajasthan state.

The radioactive waste immobilization method being adopted at Tarapur is a semi-continuous process which incorporates the oxides of high level waste from fuel reprocessing into borosilicate glassy matrices. The nominal throughput of the plant is about 25 litres of waste an hour, equivalent to the immobilization of wastes produced from reprocessing about one tonne of spent fuel a day. The plant would produce about 125 kilogrammes of glass a day, which is to be stored in stainless steel containers.

The process was developed in India and differs in many respects from the British and French techniques. The wastes are first calcined with glass fines in a process canister and then melted into a glass form. After sufficient "soaking" to help achieve homogeneity, the glass is cast into storage canisters — which can be inexpensive containers as the glass waste is more easily stored than liquids.

The storage units are stored in an engineered near-surface facility where they undergo continuous natural draught cooling to dissipate the decay heat under constant surveillance. The plant is designed to provide cooling for 20 to 30 years and incorporates facilities for establishing the characteristics of the waste products when the time comes for ultimate disposal.

The ultimate disposal medium has not yet been chosen, although current research favours a hard rock formation such as gneisses or granite.

The high-level radioactive wastes arise from the reprocessing of spent uranium fuel for the extraction of plutonium. The Atomic Energy Commission now proposes to build a chain of reprocessing plants in the country near the sites of future nuclear power stations. The proposal follows the United States' refusal to provide enriched uranium for the Tarapur nuclear power plant unless India meets conditions specified by the United States in the name of nuclear non-proliferation. Plutonium extracted by the reprocessing of spent fuel would therefore substitute for supplies of enriched uranium withheld in this way.

Sunil Saraf

CORRESPONDENCE

Creative engineers

SIR — Darnbrough (*Nature* 18 March, p.192) and Austen (*Nature* 25 March, p.284) both miss, or intentionally obfuscate, my analogy about the evolution of aeroplanes and biological evolution. Both processes lead from the simple to the highly complex. Hoyle and Wickramasinghe, not I, chose the aeroplane simile and implied that evolutionists believe that "higher life forms evolved by chance", and that a gene "might be chosen from 10^{20} nucleotide sequences of the appropriate length". This implies that large complex proteins appear without any intermediate stages.

Actually, proteins are thought to have evolved by the joining together of short polypeptides, as pointed out by me (*Molecules and Evolution*, Columbia University Press, pp. 222-229, 1960), and further explained by Rout (*Nature* 18 March, p.192). Ycas has given striking examples of this process in "periodic" proteins such as collagen and keratin (*J. molec. Evol.* 2, 17; 1972). Lengthening was probably by duplication and recombination in DNA, followed by an accumulation of point mutations.

Obviously the evolution of aeroplanes, completed in a few years, was a product of the creative genius of engineers. The point was that the Boeing 747 did not suddenly appear on the scene. In contrast, biological evolution took millions of times as long, and was impelled by slow, insensate natural phenomena. Austen says that "without a Creator, atoms and molecules, which He created, are still atoms and molecules". How could atoms and molecules exist without a Creator if He created them?

THOMAS H. JUKES

University of California,
Berkeley, California, USA

A thermal history

SIR — Many letters in recent issues of *Nature* show little unanimity about the mechanisms or rate of evolution. There may, however, be agreement that organisms do not adapt to, or evolve for, life in non-existent environments: nor do they retain capacities needed in an environment in which they no longer live.

In the course of their mainly reasonable discussion on the probable thermal history of Earth, while exposed to a steadily increasing rate of inflow of solar energy, Lovelock and Whitfield (*Nature* 296, 561; 1982) quote the observation that 50°C is the "critical upper limit for most life" as a factor supporting their thesis that the overall temperature never fluctuated greatly from today's values. That particular piece of evidence is not very cogent. Hot environments, such as hot springs and submarine spreading centres, are small and, from an evolutionary and geological standpoint, evanescent. Circumstances do not now exist that favour the evolution of organisms adapted to life at high temperatures, nor the persistence of that ability if it existed earlier.

The fossil record tells us nothing about the thermal preferences of very early organisms.

Biological potentialities should not be underrated. Enzymes vary in thermostability so, if granted several million years of undisturbed occupation of an extensive hot site, there is good reason to expect the emergence of proteins more robust than those with which we are familiar. It is not just a coincidence that the boiling point of water under today's atmospheric pressure is the limit beyond which no organisms are known to be able to live, and that very few organisms can approach that temperature. There would be no advantage in the ability to exceed that temperature — or in many organisms adapting to it. And proteins are not the only conceivable vehicles for quasi-enzyme activity. Had there been a prolonged period when the overall temperature was greater than it is now, there is no reason to assume that organisms adapted to it could not have existed. Recognition of this possibility extends the scope of discussion about the origins of life.

N. W. PIRIE

Rothamsted Experimental Station,
Harpenden, Herts, UK

Not unilaterally

SIR — Though you published an editorial (*Nature* 1 April, p.380) commenting on the three letters contained in that issue replying to the editorial of 18 February (p.542) criticizing the use of professional reputation to give credibility as it were through the backdoor, to causes like disarmament, you failed to notice the essential sophism of the points in the letters. And as this sophism is of much wider consequence than the matters addressed in these letters, affecting the whole popular movement of opinion on a subject of great importance, it deserves attention and debate.

The error is seen in the tendency to deduce from the articulated description of the horrors of nuclear war, the wrongness of, and hence the desirability to prohibit, nuclear arms. It was not felt by any of your correspondents necessary, beyond showing how awful the results of nuclear war are, to argue that disarmament has any relevance to diminishing the risks of war.

It seems to me that those who in supporting disarmament say that they are primarily concerned to achieve a wider public awareness of the consequences of nuclear war, are in fact advocating response rather than consideration. They imply by arguing how wrong nuclear wars are, that the obvious conclusion is that to prepare for the eventuality of such a war is also wrong. To argue like this is to take advantage of the association of ideas in the mind, but to evade the rational consideration of what causes war.

Although it is not my intention to argue for nuclear weapons, only to point out that the greater salience of an idea does not excuse a failure to demonstrate the validity of an argument, it is important to remember that wars are initiated in the expectation, however misconceived, that winning them is possible. Nuclear weapons as presently shared between the superpowers do not allow such misconceptions, though after unilateral disarmament one fears they could.

J. R. SKOYLES

London NW3, UK

Propaganda war

SIR — Your leading article of 1 April (p.380) headed "Professional propaganda" complains that professional groups, such as physicians who state that the threat of nuclear war currently presents the most urgent challenge to preventive medicine, do not at the same time offer alternative political and military solutions to the problems of arms control and defence. I suspect that if they were to try to do so you would censure them for making pronouncements on subjects beyond their expertise.

You would, however, be mistaken to presume that those who take the step of joining organizations such as Physicians for Social Responsibility in the United States or the Medical Campaign against Nuclear Weapons in the United Kingdom are politically ignorant. It is because they know enough to be unconvinced by assurances that nuclear weapons are necessary and for the best that they try to bring home to the public and to politicians the dangers of regarding such weapons as acceptable instruments of policy. Similar concerns also stir physicians in other countries, both East and West. At a congress held in Cambridge recently by International Physicians for the Prevention of Nuclear War (an arrogant title, perhaps, but not inaccurate) some two hundred eminent participants from 31 countries were unanimous in concluding that nuclear war — whether intended or resulting from tragic accident — would be catastrophic for any countries involved and might have dire consequences for humankind, that the medical services could have only a marginal effect in mitigating the suffering of immediate survivors, and that it was a continuing duty to bring this home to the public and to politicians in their own countries.

The congress agreed some carefully prepared statements, among them letters to President Reagan and Chairman Brezhnev which included the following outright recommendations for political action:

"We believe that action must be taken now to prevent nuclear war and to relieve the consequences of the arms race. We urge that as a first step, the nuclear powers cease all further production, testing, and deployment of nuclear weapons and their delivery systems. This should be accompanied by mutually acceptable methods of verification. We further urge that all nuclear powers unequivocally renounce the use of such weapons, and agree to prevent their introduction into any conflict. We advocate effective bilateral and multilateral negotiations on the limitation, reduction, and ultimate elimination of nuclear weapons.

We are aware that negotiations have been in progress for many years with essentially no apparent effect on the arms race. While we welcome any initiative by either side to reduce its stock of nuclear weapons, we recognize that they are likely to be eliminated only by negotiation. We must again urge that such negotiations be pursued seriously and continuously until they succeed."

J. H. HUMPHREY

London NW6, UK

Britain's resources of coal and spent uranium fuel

Robert Olby*

THE United Kingdom is often described as a favoured nation in an increasingly energy-hungry world — possessing riches of coal, and also North Sea oil and gas and an abundance of spent fuel from nuclear reactors. Before the OPEC oil price rise of 1973 it was generally accepted that dependence on coal should continue to decline in favour of oil and nuclear power. The imposition of a fuel tax on oil in 1961 marked an attempt to slow down the rate of oil substitution, but nevertheless between 1959 and 1973 oil consumption more than doubled, and by 1973, 3 per cent of total primary energy was supplied by nuclear power. The action of OPEC not only strengthened the argument for nuclear power based on its contribution to national security but within the National Coal Board it brought to the fore the already active concern to reassess the future of coal.

In the board's *Plan for Coal* (1974), investment of £1,400 million over the following decade was proposed, £8 million, of which was to be spent on an intensified programme of exploration to decide on the best sites for future development. Three years later, this expansionist mood became apparent to all with the announcement of the board's revised estimates of the United Kingdom's "physically recoverable" reserves. To a British Association audience the then-chief geologist to the board, A. Michael Clarke, declared that economics and a lack of sufficient physically recoverable reserves of coal were not the sole determinants of the choice between coal and nuclear future bulk energy supplies for Britain and Europe. Just because the era of cheap oil had taken an unnecessary slice from the life of UK coal reserves, it was not the case that in the long-term we lacked adequate reserves. To go nuclear, he declared, was not therefore inevitable.

The United Kingdom's reserves, which the World Energy Conference *Survey of Energy Resources* for the reference year 1975 (published 1977) had shown as 3.887×10^9 tonnes, were now claimed by the board to be 45×10^9 tonnes. Small wonder that this apparent dramatic change caused adverse comment from the Institution of Geological Sciences (IGS), letters in *The Times* and a tart editorial in *Nature*.

In 1975 Sir John Hill, chairman of the Atomic Energy Authority, drew attention to the potential energy source available in spent uranium fuel. With the fast reactor, he claimed, Britain's spent fuel could yield

as much as 50×10^9 tonnes of coal, more, he declared than Britain's total coal reserves which, according to the National Coal Board in 1969 were about 20×10^9 tonnes. In 1976 Sir John's colleague T.N. Marsham gave the more conservative coal equivalent of stored spent fuel as 20×10^9 tonnes, which Sir John later doubled to 40×10^9 tonnes a year. Between these two values came the Coal Board's new estimate of Britain's coal reserves at a little more than twice T.N. Marsham's figure.

rate of expansion of coal extraction of the first half of that century (800 per cent). Such a drain on the country's "life-blood" would exhaust its resources by AD 2034. On the other hand, if, as Edward Hull believed, the coal industry would be incapable of raising more than 100 million tons per annum, the nation still had eight centuries of supplies. Meanwhile he looked to science to show how we could win energy from "the light and heat that is everywhere around us". Appealing to God's provi-

During the 1970s estimates of the United Kingdom's reserves of coal and spent fuel from nuclear reactors, provided by the National Coal Board and Atomic Energy Authority respectively, have tended to steadily increase. It is tempting to infer a connection between these estimates and the efforts of the NCB and AEA to promote their claims on government finance, by each predicting as long-lived a future secure energy source as its rival. This article follows the history of the various estimates and attempts to assess their credibility.

Dwindling resource

Although the long-term future of coal as the United Kingdom's chief energy source was a matter for concern after the Second World War, the subject had long been debated. In 1789 John Williams discussed the "limited quantity of coal in Britain" in his book *Natural History of the Mineral Kingdom*, and in the 1810 edition edited by James Miller, Macnab's 1792 estimate of 360 years' supply was given. This figure related only to the coalfields of Northumberland and Durham and the extraction rates of that period. In the nineteenth century the Oxford geologist, William Buckland, urged upon the parliamentary committees of 1830 and 1835 the need to conserve coal. In 1830 the old sea-coal tax had been abolished, but twelve years later Sir Robert Peel put a tariff on all exported coal.

The subject of coal resources came before Parliament once again when the Commercial Treaty with France was debated in 1860. According to Article 11, no duty was to be levied upon coal exported to France and other nations with whom the United Kingdom was at peace. It was the discussion of this treaty which caused Edward Hull, member of the Geological Survey, to write *The Coalfields of Britain* (1860) in which he took a sanguine view. Introducing the mining limit of a depth of 4,000 feet he reckoned England and Wales had just under 59×10^9 tons, enough for about a thousand years. In the second edition of 1862 Hull included the Scottish coalfields which raised his estimate to 79.8×10^9 tons. Hull recoiled from the implications of continuing indefinitely the

dential character he assured his readers: "nor can we suppose that any part of the Creator's universe has been regulated on so short-sighted a plan, that it shall become disorganized because some of the elements necessary to its economy have failed".

The optimism of Edward Hull was matched by the stern realism of the economist and statistician, W. Stanley Jevons. In his celebrated work, *The Coal Question* (1865) he calculated the duration of British coal on the basis of Hull's estimate of reserves and the extrapolation into the future of the historic trend in coal consumption, which had been 3.4% per annum over the decade 1851–61. His conclusion was that: "Rather more than a century of our present progress would exhaust our mines to the depth of 4,000 feet . . ." Jevons was widely misunderstood as predicting exhaustion around the year 1975, whereas he was really trying to show that the nation's industrial expansion could not continue for long at the rate then current. Whilst there was still plentiful cheap coal attempts should be made to reduce the national debt, introduce a general system of education, and impose a far more general restriction on child labour. Such tasks would be more difficult to achieve when other countries reached and surpassed the output of Britain's mines, becoming more competitive in world markets (by 1967 US pithead prices were half those of Britain). Although Jevons foresaw this situation he could not, as a free-trader, recommend the imposition of an export tariff on coal. Therefore it was clear to him that the cost of coal extraction was a decisive factor in determining the

*The University of Leeds Department of Philosophy, Leeds LS2 9JT, UK.

proportion of Britain's total coal resources which would be extracted. In the future we might, indeed, be able to mine coal at greater depths than at present, but only if the price was competitive in world markets. There was no question of the physical exhaustion of our coalfields — "though we may some day have to pay dear for fuel; it will never be positively wanting", he explained in the second edition of *The Coal Question*. Here Jevons was expressing what later became a major feature of the distinction between reserves and resources, the history of which is recounted below.

Jevons drove his point home by a demonstration of the unique character of coal. In a remarkable chapter on substitutes in the third edition of his book (1906), he discounted sunlight, peat, hot springs, water and wind power, the tides, petroleum, hydroelectricity and hydrogen produced by the electrolysis of water. He found it absurd to picture the 7,308 horses or 1,000 large windmills which would be required to power a modern factory like the great Dowlais ironworks, and he pointed to the fallacy of considering electricity as a source of self-creating power. The chief targets of this critique were the optimists, Dionysius Lardner, Charles Babbage and Edward Hull.

Sir John Herschel applauded Jevons. "Such a work as yours", he wrote, "has long been wanted to dissipate *completely* the delusion which so large a majority of our countrymen labour under, of the 'inexhaustibility of our mineral resources' etc. and the 'probability amounting to certainty' that science will ere long put us in possession of a substitute for coal. . . after this let no man plead ignorance and say 'who would have thought it'." Herschel went on to describe his favourite notion of using the tides by running pipes under the sea to a distance of 1,000 feet around 1,570 miles of coastline from Berwick-on-Tweed to the Solway Firth. With a 6-foot tidal fluctuation he reckoned this vast construction would supply only one million tons of coal equivalent — one fifth of London's annual consumption or enough to keep ten Great Easterns constantly under steam.

By the summer of 1866, Jevons had become a newspaper celebrity. Gladstone wrote to him, J.S. Mill quoted him and Queen Victoria appointed a Royal Commission on coal chaired by the Duke of Argyll. In its report five years later, this commission estimated the amount of coal which "may reasonably be expected to be available for use" as 90.2×10^9 tons in ascertained coalfields, plus a possible 56.3×10^9 tons elsewhere. The resulting total recoverable resource figure of $\sim 146 \times 10^9$ tons within a depth of 4,000 feet appears to have been accepted.

The commissioners were sceptical of making projections of future demand based on historic trends; they doubted that Britain's coal would ever be *absolutely* exhausted, but the country would lose its

advantageous position in regard to coal supplies, coal imports would become the rule rather than the exception. Ominously the report concluded with the sentence: "But it may well be doubted whether the manufacturing supremacy of this kingdom can be maintained after the importation of coal has become a necessity." These views were not substantially revised by the Royal Commission of 1901, nor by anyone else.

Like their 1871 predecessors, the 1901 commissioners misunderstood Jevons. They arrived at an estimate of total available coal to 4,000 feet of nearly 142×10^9 tons, 40.7×10^9 tons being in "unproved" coalfields. The approximate agreement between the two commissions' estimates, despite the quantity of coal extracted in the interval, was due in Jevons' son, Herbert's, opinion "to the well known law that a number of unbiased errors tend to cancel one another". Nevertheless it consolidated a growing consensus that Britain had at least 140×10^9 tons of which the proportion in "proved" coalfields was increasing due to continued exploration. Conversely, the consensus was developing that as more of the best and most accessible coal was removed, the remainder would be increasingly difficult and expensive to mine. Therefore, the 1901 commission declared, the rate of increase of extraction, "will soon be checked by natural causes" so that there was "no present necessity to restrict artificially the export of coal in order to conserve it for our home supply".

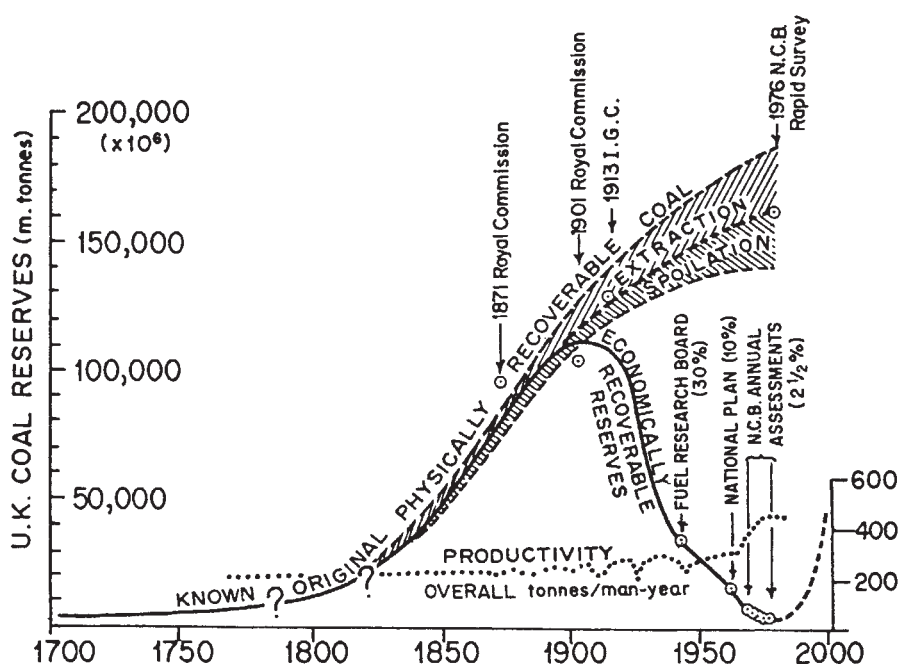
In the inter-war period, the coal industry laboured under many difficulties, the chief of which was falling demand. By the time of nationalization in 1947 it was in a parlous state, having suffered from many years of financial stringency and lack of investment. Despite extensive mecha-

nization output had not grown as it had in continental mines. This was chiefly due to the winding and undulating character of the underground lay-outs ("in-seam mining") in Britain in contrast to the straight roads which were driven through the strata on the continent ("horizon mining"). The latter layout was easier to mechanize throughout. In Britain the work of the coal-filler at the end of the conveyor belt remained unmechanized and constituted the bottle neck in the system until the 1960s.

Although Britain did not suffer so acutely after the Second World War as other European countries from coal shortages, demand was depressed by the system of allocation to consumers and there were frequent electricity cuts. Even allowing for the growth of the oil industry with Marshall Aid a long-term reorganization of the coal industry would be required to expand production. Such anticipated expansion did continue until 1958 when the threat of cheap oil began to bite into coal's traditional markets confounding the prophets of its future demand.

In such difficult circumstances mines were being closed at a rate of 40 a year, and the economically workable reserves were written off (depleted) at an alarming rate. The result is seen in the figure below.

The value of 3.87×10^9 tonnes appeared in the 1974 World Energy Conference *Survey of Energy Resources* alongside a figure for total coal reserves of 99×10^9 tonnes and almost 163×10^9 tonnes for total resources. It did not escape the notice of the conference report that the United Kingdom's recovery percentage was thus only 4 per cent in contrast to West Germany's 40 per cent. On his return from



Historic changes in estimates of UK coal reserves and productivity (tonnes per man-year). Adapted from A.M. Clarke in *Australia's Mineral Sources* (eds G.M. Philip & K.L. Williams, Sidney, 1978).

the conference Leslie Grainger, the National Coal Board's Member for Science, pointed out the absurdity of these data to the board.

A hint of what had happened can be found in the World Energy Conference *Survey* for reference year 1975. Although total reserves were now nearly 101×10^9 tonnes, economically recoverable reserves were still only 3.887×10^9 tonnes. From a footnote we learn that this figure concerned reserves at existing collieries and opencast sites (what are usually described as "operating reserves"). A second footnote tells us that the National Coal Board's "alternative" estimate for economically recoverable reserves (not restricted to those with access from existing mines and by local marginal costs) was 45×10^9 tonnes.

Resource terminology

Was this confusion merely a product of the semantics of resource estimates or of some deeper issues? Clearly there is still not a consensus over the former if only because different resources tend to be suited to different terminological rules and the needs of the smaller commercial mining companies are not the same as those of a nationalized industry. The roots of a precise terminology can be seen in Herbert Hoover's little classic, *The Principles of Mining* (1909), where the tripartite division of mineral ores into proved, probable and prospective was recommended. Four years later the staff of the Canadian Geological Survey, in their task of assessing world coal resources for the twelfth International Geological Congress, asked for returns to be classified into two groups (above and below 4,000 foot depth) and these to be subdivided into actual, probable and possible. These terms, or their equivalents, were used by mining engineers and mining geologists in the preparation of expert advice on the likely profitability of mining ventures. They expressed the level of confidence to be placed on the extent and availability of the resources in the mine, and were intended to protect the investor from exposure to fraudulent claims made on behalf of unsound commercial ventures. Therefore they referred to existing technical and economic conditions.

During the Second World War the United States government, concerned about their mineral resources, held Senate hearings on the subject at which the Bureau of Mines and the Geological Survey (USBM/USGS) proposed a new tripartite definition of reserves as measured, indicated, and inferred, each division representing defined levels of reliability, the highest being the ± 20 per cent of measured reserves. The long-term needs of such national assessments led S.G. Lasky in 1949 to make an important distinction between reserves and resources, two terms hitherto used interchangeably. Whereas reserves (measured, indicated and inferred)

were known to exist and had "aspects of usability within a practical limit of time and within a specified set of economic and technologic conditions", resources comprised "all materials in the ground, discovered or undiscovered, usable at present or not, . . . considered within the context of all factors . . . that may influence its conversion into a 'reserve', and . . . that enter into prediction or opinion as to possible future usability". This distinction was recognized in Blondel and Lasky's 1956 equation:

$$\begin{aligned} \text{Resources} &= \text{Reserves} \\ &+ \text{Marginal resources} \\ &+ \text{Submarginal resources} \\ &+ \text{Latent resources} \end{aligned}$$

It has become increasingly recognized internationally, and has influenced the World Energy Conference classification whose 1974 *Survey* definitions were:

- Total resources: "The total quantities available in the earth that may be successfully exploited and used by man within the foreseeable future."

- Total known reserves-in-place: "The corresponding fraction of resources that have been carefully measured and assessed as being exploitable in a particular nation or region under present local economic conditions using available technology."

- Recoverable reserves: "That fraction of reserves-in-place that can be recovered under the above economic and technical limits."

- Additional resources: "All other classifications with a lower degree of geologic certainty as to their existence than those indicated as known."

Of this last category the 1980 *Survey* remarked that they:

embrace all resources, in addition to proved reserves, that are of at least foreseeable economic interest. The estimates provided for additional resources reflect, if not certainty about the existence of the entire quantities reported, at least a reasonable level of confidence. Resources whose existence is entirely speculative are not included.

What, then, is the meaning to be attached to the United Kingdom's 1977 figures of (in tonnes):

Total reserves	100×10^9
Economically recoverable reserves	45
Additional resources	45
Total resources	190

The National Coal Board maintains that within 1.2 km depth and in seams over 60 cm thick there are about 190×10^9 tonnes of coal, of which some 100×10^9 tonnes are well ascertained but no figure has been given for the confidence level. These estimates are in line with the general trend of the Royal Commissions of 1871 and 1901. The figure of 45×10^9 tonnes for economically recoverable reserves shows the recognition of the fact that recovery of what is known to be in the ground has fallen from the nineteenth-century level of between 70 and 80 per cent to about 50 per

cent. This fall, due to both economic factors and to mechanization, could in Clarke's opinion be reversed under high oil prices and a vigorous programme of coal exploration. Over the next decade K. Moses thought the figure for economically recoverable reserves might well rise to 50 or even 60×10^9 tonnes. Clarke suggested the possibility of returning to the older recovery percentages which would push the figure up to 90×10^9 tonnes. Before the oil price rise of 1973, George Armstrong, then chief geologist to the National Coal Board assessed the current recovery of reserves at about 25 per cent, but he clearly anticipated a major change in the industry given increased investment. With capital, he explained, "access could be made to the best seams in the vast extensions of our existing coalfields", and we should see "for the first time since the latter part of the nineteenth century . . . an increase in the reserves of coal in Britain which can be considered as economically workable". He never denied the 200×10^9 tonnes total UK resources but considered this figure of "only academic interest" and "grossly erroneous as an estimate of economically workable reserves". It was right that under stricter rules of resource estimation the old UK figures, like the early American ones, should have been reduced. It was not that some of this coal had never existed, or had vanished since 1900, but that the onus was put on the coal industry and the nation to support the exploratory research and technical development needed to bring more and more of the resources into the reserves, and thus afford a clearer picture of the future potential of our coalfields.

Nevertheless, it should be stated that there is not yet a consensus between geologists within and without the coal industry as to how much of the coal in the UK's abandoned fields remain accessible for future extraction.

Towards atomic power

The recognition of the industrial advantage to be reaped by the nation which leads the world in the exploitation of a new source of power lay behind the public and governmental enthusiasm for exploiting the atom. Radioactivity was considered as a possible novel source of power long ago in the *Times Literary Supplement* for 17 July 1903. Radium was considered amongst those "gigantic possibilities, which are not probabilities at all yet . . .". Frederick Soddy, who had "ghosted" the article, pictured how much energy radium would yield if its rate of disintegration could be speeded up. Rutherford, the article reported, "has calculated from his own experiments and those of Curie that the energy stored up in one gramme of radium is sufficient to raise 500 tons a mile high. An ounce would therefore suffice to drive a 50 horsepower motor car at the rate of 30 miles per hour round the world". If it could be achieved with radium if could be achieved with uranium and thorium. "Our

fathers busied themselves with speculating what would become of us when the world's supply of coal was exhausted. A single step of science is needed for that problem to be answered in a manner beyond the dreams even of the scientific novelist."

We all know that these prophetic words were fulfilled (for uranium not radium) when in 1939 Lise Meitner and Otto Frisch interpreted the experimental results of nuclear fission; and Halban, Joliot and Kowarski, Fermi and others discussed the possibility that it might be accompanied by a chain reaction. That year was not only the *annus mirabilis* of nuclear fission, it also marked the first attempt to calculate the potential electricity output of a power station run from a "uranium boiler". S. Flügge calculated in *Die Naturwissenschaften* that complete fission of one cubic metre of uranium oxide powder would yield 7×10^{10} kWh of electricity, sufficient to replace the entire output of Germany's power stations running on middle-German brown coal for 11 years.

Comparisons such as Flügge's have been made repeatedly since 1939 in order to give some conception of the contrast in the concentration of energy in fissionable fuels as compared with fossil fuels. Hitherto the largest yield from a nuclear transformation had been the 22.2 MeV from the conversion of lithium (^6Li) into alpha particles (^4He). But when M.C. Henderson measured the energy yield from uranium in 1939 it was a staggering 175 meV (theoretical value, 198 MeV). (The beta decay of radioactive fission products subsequent to fission of ^{235}U is 21 MeV, and has to be subtracted from the 198 MeV.) The table below illustrates the million-fold increase in yield per atom when we go from combustion to fission.

Theoretical energy yields per atom

^{12}C	Combustion	4.17 eV
^{226}Ra	Fission	4.79 MeV
^{235}U	Fission	198 MeV

When we compare uranium and carbon weight for weight the ratio is approximately 3 million to one. But H.D. Smyth in his report on the Manhattan project considered what the fission/combustion ratio would be if by fission all the matter of the atom was converted into energy. The equation relating matter and energy ($E = mc^2$), he wrote,

... shows that one kilogram (2.2 lbs) of matter, if converted entirely into energy, would give 25 billion kWh of energy. This is equal to the energy that would be generated by the total electric power industry in the United States (as of 1939) running for approximately two months. Compare this fantastic figure with the 85 kWh of heat energy which may be produced by burning an equal amount of coal.

Smyth knew then, as we know now, that less than 0.1 per cent of the matter in uranium can be converted into energy by fission (to be precise it is 0.091 per cent). His calculation showing a ratio of 3×10^9

to one was therefore knowingly a thousand-fold exaggeration. He also knew what Flügge did not know, namely that only 1 in 139 (usually stated as 1 in 140) of the atoms in uranium is readily fissionable with slow neutrons, that is, those of the isotope ^{235}U .

As later experience showed, the utilization of uranium by Britain's Magnox reactors is about 0.5 per cent (of which ^{235}U contributes about 0.35 per cent and the rest comes from the isotopes of plutonium and other elements). Taking a very high grade uranium ore with 1% uranium content, one fifth of which is recovered in extraction, and comparing it with a second-grade, bituminous coal (50% carbon) and some 30% of stone chips and dust our 'from the ground' ratio for Magnox reactors would be:

$$\frac{3 \times 10^6}{100 \times 200} \times \frac{1}{5} \times \frac{100}{35} \approx 86.1$$

Research carried out in the United States during the Second World War made uranium look much more attractive than the 21,000:1 ratio suggested, for the elements heavier than uranium, elements 93 and 94 in the Periodic Table, were discovered as products of neutron capture by the abundant isotope of uranium, ^{238}U . By 1941, Glenn Seaborg and co-workers had isolated element 94 and called it plutonium, ^{239}Pu . As expected from theoretical considerations it proved fissile together with ^{240}Pu and ^{241}Pu by both thermal and fast neutrons.

Estimates of world uranium supplies won a new significance in December 1951 when the US experimental breeder reactor (EBR 1) yielded electricity to the grid and it had been shown meanwhile that the number of atoms of the abundant ^{238}U converted to fissile ^{239}Pu exceeded the number of ^{235}U atoms fissioned. That meant that more fuel was 'bred' than was consumed.

The principle of the breeder reactor having been experimentally established, energy forecasters waxed confident about a world energy future based on uranium. Palmer Cosslett Putnam in his report to the US Atomic Energy Commission in 1953 included all low-grade sources of uranium in his calculation of recoverable world reserves at 25 million tonnes, with an energy yield of 1,716Q, one third of which — 575Q — it would be economical to 'burn'. As he was contracted by the commission to estimate the maximum possible contribution which nuclear power could make, it is hardly surprising that this estimate of uranium reserves was wildly exaggerated. Today figures in the range 6.6 to 14.5 million tons have been dismissed as misleadingly fanciful. Total world reserves are still put at less than 2 million tons, plus additional resources of little above 1.5 million tons. When Hans Thirring reworked Putnam's uranium estimates in 1958 he neglected or rejected the one third utilization which Putnam made to allow

for losses during recycling of the fuel in the breeder, with the result that Thirring's value for uranium reserves was 2,000Q. "The date of depletion of the energy capital of our Earth", he concluded in *Energy for Man*, "is therefore prolonged by at least a millenium . . . Thus the consequence of Hahn's great discovery will free us from the fear that our energy resources will run out within quite a short period of history."

US coal resources

A more surprising aspect of Putnam's report was his reworking of the estimates of US coal resources. To the 1944 figure of $3,100 \times 10^9$ tonnes (total resources) he applied a series of correction factors which reduced it by nine tenths. As Schurr and Netschert have well said Putnam's resulting 3.1Q's worth of recoverable US coal reserves "is based on limiting criteria that have no consistent basis; some of them are wholly arbitrary and some of them have no relation to coal actually in stock".

Nuclear scientists have to their credit that they have in the past cautioned against the acceptance of wildly enthusiastic claims for nuclear power. Such euphoric remarks included the "pill-in-a-pail" suggestion made in the US Congress in 1946. All you needed to heat your house for a year, according to this idea, was a small pill of uranium in a pail of water. Other calculations suggested that a piece of uranium the size of an egg would propel the *Queen Mary* from New York to Europe and back, whilst other pundits claimed that one the size of a pea placed under the doorstep would suffice to heat a house for its entire lifetime.

At the end of the Second World War, experience with the Hanford reactors had shown that very little of the 0.7 per cent ^{235}U in the fuel was 'burnt' before 'poisoning' seriously reduced fission. This led Oppenheimer in 1947 to advise that it did

not appear hopeful to use natural uranium directly as an adequate source of fuel for atomic power. The reactivity of systems based on natural uranium is low; even the fraction of ^{235}U which can be consumed without replenishing the fuel is small. Because of this the raw material requirements, if such reactors are to play an important part in power economy, are economically prohibitive. It is true that one could re-enrich this natural fuel, but the power expenditure involved in such isotope separation by any present methods makes this likewise prohibitive.

Although the General Advisory Committee to the Atomic Energy Commission discussed drafts of Oppenheimer's memorandum, it was not published for fear of undermining congressional support for the commission. Oppenheimer's ground for pessimism over natural uranium reactors has, as we know, proved exaggerated. Britain's Magnox reactors do 'burn up' some 0.5 per cent of the entire fuel since they fission about half

the ^{235}U and some of the plutonium they produce. Nevertheless reactors of this type are no longer built, precisely because they use so little of the fuel.

Considerable caution was also expressed at the first nuclear power session of the World Power Conference when the subject was included in this organization's Fuel Economy Conference at The Hague in 1947. Charles Thomas's calculation of the cost of nuclear electricity (nuclear 0.8 cents per Kwh, coal 0.65 cents) for the UN Atomic Energy Commission in 1946 was attacked by Con. Edison's research engineer, Ward F. Davidson and others. Both Sir John Cockcroft and L. Kowarski were cautious, the latter being decidedly caustic about the fast reactor. The tone was again cautious at the full session of the World Power Conference in 1951. Cockcroft said it was "like gazing into the crystal ball to ask us at this stage what operating and capital costs are going to be". At the Atoms for Peace Conference of 1955 he was looking to this "second decade" (the 1970s) for the uranium/coal ratio to be raised by the breeder principle to at least a million to one.

It was in 1951 that work on the fast reactor had begun in the United Kingdom, stimulated by concern over uranium supplies. Eleven years later, the Dounreay fast reactor (DFR) began generating electricity. This 60 MW(t) station was fueled with uranium, 75 per cent enriched ^{235}U , not plutonium. It was run on full power for five years before being treated as a test bed for plutonium fuels. It has been succeeded by the prototype fast reactor (PFR) using mixed oxide fuel. By 1973 construction was completed and the following year saw operation at low power.

It is surely no accident that the potential world role of nuclear energy was being stressed in the United Kingdom from 1973 onwards based upon the promise of the fast reactor. Its high fuel utilization, claimed Sir John Hill in the magazine *Atom*, "will enable the vast reserves of very low-grade uranium to be utilized without significantly increasing the overall cost of electricity". World reserves of coal, he admitted, were enormous, but "we are consuming the most accessible and most convenient fuels (gas and oil) fastest and we are depleting most rapidly those reserves which are most conveniently situated geographically, economically and politically". Coal, though abundant, was expensive to mine and transport. Moreover the electricity industry needed flexibility to cope with shortages such as strikes and embargoes could create. Though improvements in UK coal production were impressive, the industry remained in Sir John Hill's words "essentially a labour intensive industry despite increasing use of machinery and, if the men employed are to maintain a standard of living compatible with the task they are being asked to undertake, the labour costs of the industry will rise and this must be reflected in the cost". The

context of these words was provided by the coal price rise of 40 per cent over eighteen months in 1969–70 and the picketing of power stations by mineworkers in 1972. More was to follow — the 140 per cent price rise of coal in 1975, the 400 per cent price rise of oil in 1973–74 and the miners' strike of 1974 which involved a four-month dispute in the industry.

Conclusion

Statistics on fuel reserves do not constitute cold hard facts, they are judgements about the probable outcome of a series of events in space and time. One might even claim that they do not so much form government policy as result from such policy. Since the introduction of oil-fired power stations in the 1950s the policy of building a multi-fuel base to electricity generation with nuclear base-load has been followed by both political parties. Whereas governments have in the past given some protection to coal against oil there has been from 1955 onwards a consistent policy of favouring nuclear power, a policy shared by the Central Electricity Generating Board whose aim it is to increase dependence upon nuclear-generated electricity for base load, and thus to diminish the impact of higher coal prices and the threat posed by the withdrawal of labour in the coal industry. When the all-party Commons Committee on Energy delivered its report in February 1981, a *Daily Telegraph* editorial described the recommended "cut-back in nuclear power station construction" as "naive, petty-minded, and ill-informed". The editorial continued:

It is no doubt bad luck for its authors that the report appeared just as an indefinite national coal strike was threatened. But we, on the other hand, should be thankful for the timing. It enables us to see how absurd it would be not to invest massively in nuclear power at a time when coal supplies are under constant threat, and, more important still, when the Russians are reaching out greedy claws towards the Middle East oil wells.

Clearly the formulation of energy policy is a highly political process. The decisions reached can affect the estimates of reserves themselves. Investment in machinery, prospecting, methods of combustion and of fission, all influence what proportion of the total resource may be considered as economically extractable reserves. Thus fluidized-bed-combustion can be achieved profitably with low grade coal of ash content up to 55 per cent instead of the maximum 30 per cent for conventional combustion, thus bringing high-ash coal deposits into the reserves. Similarly high-sulphur coals whose combustion causes acid rain can be burnt without liberation of sulphur dioxide in the fluidized bed. Uranium reserves, which at present are trivial by comparison with coal, are increased in terms of their energy yield some sixty-fold with the fast reactor. This, in turn, will allow lower-grade ores to be brought from the resource category and

FOR an excellent critical discussion of resource nomenclature the reader is referred to G. B. Fettweis, *World Coal Resources: Methods of Assessment and Results* (Elsevier, Amsterdam, 1979). Also invaluable are: The Royal Society's submission to the Commission on Energy and the Environment, *Environmental Aspects of Increased Coal Usage in the U.K.* (February 1980) and the WATT Committee on Energy, *Assessment of Energy Resources, Committee Report 9* (1981) and the report of the Commission on Energy and the Environment, *Coal and the Environment* (1982). The proceedings of the November 1981 British Nuclear Energy Society conference on the fast reactor fuel cycle should be published soon. There the "reasonably assured resources" of uranium are given as 2.6 million tonnes.

added to the economically exploitable reserves. There is a sense, then, in which, as Clarke has remarked, an investment decision which significantly favours one energy industry at the expense of the others, is likely to be self-fulfilling, and may thus prove to have been economically and financially correct.

The claimed potentials of both coal and spent uranium fuel reserves raise sensitive issues — fears of increased seismic activity, water table changes and water contamination, in the case of coal mining; concerns about reprocessing and long-term storage, and doubts about the commercial performance of the fast reactor, in the case of spent uranium fuel. If the Coal Board's estimate of 45×10^9 tonnes looks like a "giant extrapolation" to quote a *Nature* editorial, how much better are the claims made on behalf of spent uranium fuel? True it has not to be mined, but it will have to be reprocessed many times before even half its energy content has been liberated.

This analysis of energy resource claims does not suggest intentional rivalry by the Atomic Energy Authority and the National Coal Board. Their presentation about the same time was fortuitous. That they both represent responses at the political level in the campaign for government support seems evident. Also clear from this analysis is the fact that the modern distinction between resources and reserves, though dating from 1949, is still in the process of winning general acceptance. Reluctance to adopt it might well be due not only to tradition but also to the tendency which strict definitions of economic reserves have to belittle the long-term potential of an industry's resources. If decision-makers have some appreciation of the factors involved in the process of converting resources into reserves then the application of this fundamental distinction of nomenclature will be more consistently adopted.

I acknowledge the assistance of Mr A. Michael Clarke, Dr Leslie Grainger and Professor E. H. Francis; Mr Keith Bowes kindly commented on an earlier draft of the section of the paper given to nuclear energy. The extract from Oppenheimer's memorandum to the AEC Advisory Committee is cited with the permission of the US Department of Energy. □

NEWS AND VIEWS

Not quite full circle? — non-racemic amino acids in the Murchison meteorite

from C. T. Pillinger

In this issue of *Nature* (p.837), Michael H. Engel and Bartholomew Nagy report finding partially racemized glutamic acid, aspartic acid, proline, leucine and alanine in the Murchison meteorite. They believe the amino acids to be indigenous to the sample, even though full racemization is usually accepted as the indicator that such compounds were present before the meteorite's arrival on Earth. Nevertheless, the authors present a good case and, to their credit, make no attempt to invoke extra-terrestrial biology as an obvious and sensational interpretation of the results. Engel and Nagy only speculate that some type of stereoselective synthetic or decomposition reactions could have taken place on the meteorite's parent body. Less responsible scientists will not be so cautious or await proof that some unusual form of terrestrial contamination was not responsible for the differences between the latest results and samples previously studied.

This publication could be said (unkindly by some) to set back the study of carbonaceous substances in meteorites by some twenty years since it was in 1961 that Nagy, writing with W.G. Meinschein and D.J. Hennessy¹, presented the results that sparked off a decade of controversy. Using organic mass spectrometric techniques which were just finding widespread use in the petroleum industry, the group had discovered aliphatic and aromatic hydrocarbons in the Orgueil meteorite. At the time such precise chemical structures were considered as 'chemical fossils' or 'biological markers' when found in ancient sedimentary rocks and petroleum. The analogy seemed simple — carbonaceous chondrites must be the extra-terrestrial counterparts of such materials. In voicing this conclusion, Nagy and his co-workers were following the suspicions of previous generations who, beginning with Berzelius², questioned whether the results of meteorite analyses were providing 'a hint concerning the presence of organic structures in other planetary bodies'.

After 1961, a period of vigorous activity ensued, during which many of the classes of compounds common in the Earth's sedimentary regime were located in one or other of a few known carbonaceous chondrites (see ref.3 for a review). Similarly,

electron microscopic searches for microfossils were initiated and became the subject of acrimonious debate — another controversial area that has just been resurrected (H.D. Pflug, Public Lecture, Cardiff, 26 November 1981).

All the investigations faced the problem that the samples being studied were not fresh; carbonaceous chondrites are extremely porous and friable and some of the most interesting over 100 years old. Who could know what dubious fate had befallen them at the hands of loving collectors and curators? Then, in 1969, the necessary samples were provided. First, on 8 February, a large shower of material landed near Pueblito de Allende, Mexico, almost on the doorstep of the expectant Lunar Receiving Laboratory at Houston. Within possibly hours and certainly two or three days, large samples were being carefully stored free from contamination. Yet despite these precautions, appreciable quantities (1 p.p.m.) of 'biological' hydrocarbon and fatty acids were found⁴ in exterior, but not interior, portions of the meteorite, presumably having accrued during the brief passage through the atmosphere and sojourn on the ground. The lesson was salutary.

But Allende was still not the right meteorite: that arrived later the same year near Murchison, Victoria, Australia. Some of the first results reported for the sample concerned amino acid analyses⁵. By this time gas chromatographic techniques had progressed considerably and it was possible to separate amino acid enantiomers, appropriately converted to diastereomeric derivatives, on optically inactive stationary phases. Kvenvolden *et al.* found that the asymmetric amino acids in Murchison were fully racemized, that is, equimolar mixtures of the possible isomeric forms. Since amino acids from active biological systems are specifically *laevo*-rotatory or L-form, the authors were able to exclude the possibility that terrestrial organisms had contaminated their sample. Moreover, the distribution of structures was different from that encountered in biological proteins. The fact that simple molecules such as glycine, α -amino butyric acid and α -alanine were most prolific convinced Kvenvolden *et al.* that they had encountered extraterrestrial amino acids of abiogenic origin.

The Engel and Nagy paper (see p.837) is almost entirely given over to arguments supporting their conclusion that the par-

tially racemized amino acids they find in Murchison are indigenous to the sample. Briefly, these include the observation that the least racemized acids are the most difficult to extract, and the absence or only trace occurrence of common protein amino acids tyrosine, methionine and phenylalanine and the fingerprint indicators serine and threonine. Appropriate controls are offered to demonstrate that the extraction process did not cause the partial racemization or obliterate the evidence of terrestrial contamination. A selected ion-monitoring technique is used to demonstrate that the apparent imbalance in stereoisomer concentrations is not due to some unsuspected co-eluting contaminant. Finally, corroboration of the results was solicited by circulating the samples to colleagues working in the same research area. One less than satisfactory statement is the suggestion that the present study 'may not contradict previous reports, because a large interior sample was analysed'. Surely small samples would be expected to show inhomogeneities in distribution, not the large one.

Further examinations of amino acids from clean carbonaceous meteorites would indeed be beneficial. What is needed is an absolute indigeneity criterion, independent of stereoisomerism. Not long ago⁶, it was recognized that the total amino acids isolated from Murchison were isotopically very distinct from their terrestrial counterparts and polar extractables from several carbonaceous chondrites are enriched in deuterium⁷. Perhaps the intramolecular isotopic composition of individual enantiomers will provide the answer; not an easy measurement to perform but feasible in the long term.

Oh that Engel and Nagy had found excesses of D-amino acids, then one could accept the data or suggest some grudge-bearing hoaxer had spiked the sample! But that would be even more difficult to explain. After all, we are looking for a source of abiogenically produced L-amino acids to set biochemistry on its discriminatory way. □

1. Nagy, B., Meinschein, W.G. & Hennessy, D.J. *Ann. N.Y. Acad. Sci.* 93, 25 (1961).
2. Berzelius *Ann. phys. Chem.* 33, 113 (1834).
3. Hayes, J.M. *Geochim. cosmochim. Acta* 31, 1395 (1967).
4. Han, J., Simoneit, B.R., Burlingame, A.L. & Calvin, M. *Nature* 222, 364 (1969).
5. Kvenvolden, K. *et al. Nature* 228, 923 (1970).
6. Chang, S. & Mack, R. *Lunar planet. Sci.* 9, 157 (1978).
7. Becker, R.H. & Epstein, S. *Geochim. cosmochim. Acta* 46, 97 (1982).

C. T. Pillinger is in the Department of Earth Sciences, University of Cambridge, Cambridge CB2 3EQ.

Mutualistic interactions among species

from Robert M. May

MUTUALISTIC interactions between species have traditionally received significantly less attention from ecologists than have competitive and predator-prey interactions. Recent years have, however, seen growing awareness of this fact, and, in my opinion, empirical and theoretical studies of mutualistic associations are likely to be one of the growth industries of the 1980s. This was one of the themes of the Special Symposium on "Mutualism: New Ecological Theories" held under the auspices of the AAAS at its Annual General Meeting in Washington, DC, in early January.

In opening the Symposium, Boucher observed that mutualism has been cast in the role of the invisible Holy Ghost among the trio of basic types of interaction (see also the review by Boucher *et al.*¹). In elementary biology texts, the contending claims of competitive and of predator-prey relationships as regulators of population density are usually discussed, while mutualism is dismissed descriptively with some colourful pictures and engaging anecdotes. Most of the current generation of ecology texts compound this disparity: competition and predator-prey typically each have one or more chapters to themselves, replete with discussions of theoretical models and field and laboratory experiments; mutualism is generally treated in a perfunctory way.

As I have suggested elsewhere², there may be two main reasons for this relative neglect of mutualistic interactions. On the theoretical side, the simple mathematical models developed in the 1920s and 1930s by Lotka, Volterra, Kostitzin and others do capture some of the essential features of competitive and predator-prey dynamics, but they lead to nonsense for mutualism. On the empirical side, the conspicuous mutualistic associations tend to be tropical ones, whereas until relatively recently the bulk of ecological research was done in temperate regions (as a glance at, for example, the cumulative index in the *Journal of Animal Ecology*³ will show); temperate zone mutualisms certainly exist, but — as we shall see below — they can be more subtle than tropical ones.

Models of mutualistic associations

The non-linear Lotka-Volterra equations describe the interactions between two species in such a way that the per capita birth rate of species A depends linearly on the population densities of species A and of species B. In mathematical terms, this means the 'isoclines', which partition the dynamical behaviour of the populations A and B into various qualitatively distinct regions, are straight lines. For competing species, the results can be victory for species A, or for species B, or

coexistence; more sophisticated models give broadly similar behaviour. For predator-prey, the outcome is a propensity to cyclic oscillations, which again gives the key to the behaviour of more refined models. But for mutualisms that are sufficiently strong, these simple models lead to both populations undergoing unbounded exponential growth, in an orgy of reciprocal benefaction.

As reviewed at the Symposium by Post, a variety of recent analyses have sought to remedy this defect^{2,4-6}. In essence, all these studies observe that the benefit conferred by one species on another will eventually saturate to some limiting value, set by resource or other constraints. The resulting curvilinear isoclines must intersect, giving some finite population level for the mutualists. At the Symposium, Culver⁷ described a specific instance in which the limits to a mutualism (between ants and some spring herbs whose seeds the ants disperse) are apparently set by rodent seed-predators. In many of these models, the mutualistic equilibrium is one where, although both species are more abundant than they would be in the absence of mutualistic effects, both are more dynamically fragile — more prone to fluctuate and taking longer to recover from a given perturbation. On the other hand, as discussed by DeAngelis *et al.*⁸, in a patchy environment the mutualistic association between plants and seed-dispersers can act as a stabilizing influence.

Another complication arises when one (or both) of the mutualistic species comes to be 'obligate', in the sense that the population cannot persist in the absence of its partner^{2,4,9}. In this circumstance, there can be two alternative stable states, one with both species present in numbers enhanced by their association and another in which the obligate species is absent (or both are absent if both are obligate). These two alternative states are divided by a 'watershed' of population values, and environmental disturbances can carry the system from one state to the other; in practice, the more likely excursion will be from the state with both species present to the state in which one or both species are absent. An example of such a situation is given by Soberon and Del Rio⁹, whose model for the association between plants and highly specific pollinators shows that the conditions which optimize the production of nectar tend to diminish the stability of the association.

Elaborating a general notion introduced by Levine¹⁰, Vandermeer described how

effectively mutualistic interactions can arise indirectly in complex food webs¹¹. Suppose species A and B appear to compete for prey species C and D, with species B concentrating mainly on resource D. Suppose further that species C is competitively superior to D and would eliminate D were it not for the effects of the predatory species A. It now can happen that removal of species A will lead to the prey species C displacing D, thus decreasing the population of B which is sustained largely by D. Thus the apparently competing species A and B are, in fact, mutualistically related, by virtue of these 'indirect' interactions. As Vandermeer showed, when realistic time lags are added on top of these indirect effects, the dynamics of mutualistic associations can become very complex indeed. There is clearly much room for further theoretical work.

Empirical studies of mutualism

Many beautiful studies have recently been carried out, analysing the energetic costs and benefits involved in the association between nectar-producing plants and their pollinators. The plant-pollinator systems are mainly tropical, and often obligate¹²; the pollinators embrace butterflies^{13,14}, hummingbirds and sunbirds¹⁵⁻¹⁷, bats^{18,19}, bees^{20,21}, lemurs and marsupials²². Likewise, recent work on some tropical associations between plants and seed-dispersers has not only analysed costs and benefits, but has shown the long-term dangers inherent in obligate mutualisms. Thus, unless we intervene, the *Calvaria* tree looks like following to extinction the dodo that once dispersed its seeds²³. More generally, the neotropical flora may still be changing in response to the disappearance of many large mammals from South America²⁴. Fascinating though these studies are, here we will concentrate on mutualisms involving plants, ant-tended insects and ants.

The tropics afford many striking examples of obligate ant-plant associations, in which plants provide both food and housing (in special 'apartment-like' physical structures at the base of thorns or elsewhere) for ants. In temperate regions, associations in which plants offer food (but not housing) exist, albeit less conspicuously.

Inouye and Taylor²⁵, for example, have shown that Aspen sunflowers possess extrafloral nectaries that attract ants. By excluding ants from some plants, and by other manipulative experiments in the field, these authors have demonstrated that the ants confer reciprocal benefits on their host plants, by providing a degree of protection against herbivorous predators. In a similar set of carefully designed field

Robert M. May is Class of 1897 Professor of Zoology at Princeton University, New Jersey 08540.

experiments, Tilman²⁶ has shown that the extra-floral nectaries on black cherry trees attract a species of ant, which in turn helps check the depredations of tent caterpillars. A variation on this theme is provided by Messina's work²⁷ on the ant species attracted to goldenrod by the nitrogen-rich 'honeydew' they can get from tending membracids (tree-hoppers) that are feeding on the plant; as observed by Bristow²⁸, in such circumstances the phytophagous insects are a kind of portable extra-floral nectary. Here Messina's manipulative experiments again show that the ants help to exclude other herbivores from the plant.

The above studies carefully document the costs and benefits of the ant-plant association. On the other hand, they typically do not wrestle with the complexities of food web structure, often lumping different ant species together as 'ants'. The complications that can arise in unravelling the full community structure of an ant-homopteran-plant association is well illustrated by the pioneering work of Addicott²⁹: at his study site, fireweed supported four species of aphids, of which three were tended by a total of ten species of ants. This detailed investigation of community structure is, however, largely descriptive.

Of particular interest, therefore, are some recent studies which combine manipulative and analytical techniques with full dissection of the structure of the community. One such study, involving carrion-burying beetles and hitchhiking mites, was described at the Symposium by D.S. Wilson; a full report of this complex story will be given another time. Another such study involves — in its essentials — a

membracid and an aphid species that feed on ironweed in New Jersey. These homopterans are tended by two ant species, a *Myrmica* and a *Tapinoma*. Using controlled experiments in the field and in a greenhouse, Bristow²⁸ has elucidated the detailed structure of this community, along with the costs and the benefits of the association to both ants and homopterans. Under favourable conditions, there is a strong tendency towards specialization by the ant species: *Myrmica* prefers membracids, and *Tapinoma* prefers aphids. This tendency is, however, far short of an obligate dependence: the preferences are not exclusive, and either species of ant will tend either species of homopteran. Bristow²⁸ goes on to observe that in temperate zone systems, absolute dependence is often too costly. Winter forces a prolonged break in any association, in time and perhaps also in space; species exploiting ephemeral hosts must

find new hosts at least once each spring, and possibly more often. Ants cannot rely on their highly localized foraging area being recolonized by their preferred honeydew producer. By the same token, homopterans may find themselves abandoned if the ants move for some reason.

Bristow²⁸ concludes by suggesting a new paradigm for mutualistic associations outside the tropics:

"The interesting point is not that temperate zone mutualisms are rarely obligate, but rather that by relying on simple environmental cues and behavioural mechanisms a high degree of fine tuning may be achieved with almost no loss of flexibility. This suggests that facultative mutualisms may deserve a closer look than they have received; their contribution to community or guild structure may not be as negligible as has been previously presumed." □

Type I supernova — thermonuclear versus collapse

from Stirling A. Colgate and Albert G. Petschek

In a recent issue of *Nature*, Canal, Isern and Labay¹ discussed the rather esoteric problem of the evolution of a type I pre-supernova core — a white dwarf star where a eutectic phase separation of carbon and oxygen is presumed to take place before explosion. This would seem to be a singularly esoteric and even abstruse issue, but such are the labyrinthian relationships in science that this topic is central to a major issue in astrophysics — the mechanism of supernovae. It depends in turn on the most mundane property of metals — the alchemy of alloys and the long-range order functions of the fluid-solid transition, a controversial topic in theoretical physics. The same problem arises in the analysis of ordinary materials as well as of γ bursts.

A supernova — the explosion of a star — is the most dramatic real-time event that we observe in the Universe. Presumably, most of the elements necessary for life are generated and dispersed within the Galaxy by such an event. Presumably also, the most exotic energy and state of matter, cosmic rays and neutron stars, are formed in a proportion of such events. Yet there is no universal agreement as to the mechanism(s). The large star, ~ 10 solar masses ($M_{\odot}$), almost certainly makes a type II supernova, leaving a neutron star remnant. The collapse and subsequent explosion is currently modelled by the elastic bounce energy of the forming neutron star core^{2,3}. In the past, neutrino emission was thought to cause the explosion that accompanies collapse, but now neutrinos are believed to hinder the

propagation of a mass-ejecting shock. The theoretical outcome, that is, explosion or collapse to a black hole, is still dependent upon details. The need to resolve this unhappy uncertainty has stimulated considerations of possible different neutrino-matter interactions, such as axions⁴ and Majorana coupling⁵. Violent neutron star convection has also been considered by Colgate and Petschek⁶, and refuted by Lattimer and Mazurek⁷.

Type I supernovae are believed to be older, less massive stars that ignite a thermonuclear reaction in the core and are pushed over the edge of degeneracy stability — the Chandrasekhar limit — by a combination of slow accretion from a binary and evolution. If a carbon core ignites early at the low density of $2\text{--}5 \times 10^9 \text{ g cm}^{-3}$, then the star explodes from a thermonuclear deflagration and no remnant is left. However, if an oxygen core thermonuclearly ignites at the higher density of $\sim 2 \times 10^{10} \text{ g cm}^{-3}$, then rapid electron capture decreases the degeneracy pressure faster than thermonuclear deflagration drives expansion, and collapse to a neutron star occurs⁸. We are then left with the same difficulty — how does a type II supernova work? Nevertheless, we can predict that a carbon deflagration disintegrates the star and ejects all the matter ($1.41 M_{\odot}$), mostly as iron, into the Galaxy. An oxygen deflagration and collapse would eject

Stirling A. Colgate and Albert G. Petschek are at the Los Alamos National Laboratory, Los Alamos, New Mexico 87545.

1. Boucher, D.H., James, S. & Keeler, K. *Am. Rev. ecol. Syst.* 13 (in the press).
2. May, R.M. (ed.) *Theoretical Ecology: Principles and Applications*, Ch.5 (Blackwell, Oxford, 1976).
3. Taylor, L.R. & Elliott, J.M. *J. Anim. Ecol.* 50, 951 (1981).
4. Vandermeer, J.H. & Boucher, D.H. *J. theor. Biol.* 74, 549 (1978).
5. Goh, B.S. *Am. Nat.* 113, 261 (1979).
6. Travis, C.C. & Post, W.M. *J. theor. Biol.* 78, 553-571 (1979).
7. Heithaus, E.R., Culver, D.C. & Beattie, A.J. *Am. Nat.* 116, 347 (1980).
8. DeAngelis, D.L., Travis, C.C. & Post, W.M. *Theor. Pop. Biol.* 16, 107 (1979).
9. Soberon, J.M. & Del Rio, C.M. *J. theor. Biol.* 91, 363 (1981).
10. Levine, S.H. *Am. Nat.* 110, 903 (1976).
11. Vandermeer, J.H. *Am. Nat.* 116, 441 (1980).
12. Kodric-Brown, A. & Brown, J.H. *am. Zool.* 19, 1115 (1979).
13. Benson, W.W., Brown, K.S. & Gilbert, L.E. *Evolution* 29, 183 (1976).
14. Gilbert, L.E. in *Coevolution of Animals and Plants* (eds Gilbert, L.E. & Raven, P.H.) 210 (University of Texas Press, Austin, 1975).
15. Gill, F.B. & Wolf, L.L. *Am. Nat.* 109, 491 (1975).
16. Stiles, F.G. *Ecology* 56, 285 (1975).
17. Carpenter, F.L. *Am. Zool.* 19, 1105 (1979).
18. Howell, D.J. *Am. Nat.* 114, 23 (1979).
19. Howell, D.J. & Hartl, D.L. *Am. Nat.* 115, 696 (1980).
20. Heinrich, B. *Evolution* 29, 325 (1975).
21. Janzen, D.H. *Science* 171, 203 (1971).
22. Sussman, R.W. & Raven, P.H. *Science* 200, 731-736 (1978).
23. Temple, S.A. *Science* 197, 885-886 (1977).
24. Janzen, D.H. & Martin, P.S. *Science* 215, 19-27 (1982).
25. Inouye, D.W. & Taylor, O.R. *Ecology* 60, 1-7 (1979).
26. Tilman, D. *Ecology* 59, 686-692 (1978).
27. Messina, F.J. *Ecology* 62, 1433-1440 (1981).
28. Bristow, C.M. thesis, Princeton Univ. (1981).
29. Addicott, J.F. *Can. J. Zool.* 57, 558-569 (1979).

about one-fifth as much. This latter case is more comfortable for galactic theorists but comfort is not everything⁹. The light from the supernova comes from the subsequent β decay energy of ^{56}Ni that decays to stable iron ($^{56}\text{Ni} \xrightarrow{6.6 \text{ day}} ^{56}\text{Co} \xrightarrow{77 \text{ day}} ^{56}\text{Fe}$). The classic observed optical decay half life of 56 days of supernovae I¹⁰ is powered by the radioactive decay positrons from ^{56}Co . An increasing proportion of the β^+ can escape and simulate a 56-day exponential provided the ejected mass is relatively small. The larger ejected mass resulting from carbon deflagration models requires a critically timed IR emission¹¹ if agreement with observation is to be obtained. The positrons, which were once observed at the galactic centre and now seem to have disappeared^{12,13}, would have had the correct flux if the source were a supernova type I involving collapse with the escape of 10 per cent of the positrons.

Helium burning in the pre-supernova white dwarf star leads to roughly a 50:50 mixture of fluid carbon and oxygen at a density of 10^8 – 10^9 g cm^{-3} . A eutectic exists in the phase diagram (66.8 per cent ^{12}C and 33.2 per cent ^{16}O), as in a lead–tin solder. On energetic grounds, the smaller atom or ion prefers to fill the interstices of the lattice of the larger one and this lowers the melting point. The excess component, ^{16}O , precipitates out as ^{16}O snowflakes or perhaps as glass-like spherules and falls to the centre of the star. A further separation may occur after remixing of the remaining fluid due to a gravitationally induced thermohaline instability. Depending on the degree of separation of the core, an oxygen-rich core in the white dwarf will lead to a collapse, not a thermonuclear disintegration.

The energetics of the liquid–solid transition of matter is modelled in extensive Monte Carlo calculations that lead to melting point predictions¹⁴. Crystallization, on the other hand, requires additional physics for the transition to take place and has not yet been modelled numerically. The possibility of a neutron star quake leading to a neutron star volcano and subsequently a γ burst depends on the same physics of the melting transition of neutron star crust matter¹⁵.

Struggles among the dunes

from Peter D. Moore

THE manner in which one plant paves the way for another in the course of vegetation succession may sometimes have its ironic aspects. So often, a colonizing plant species nullifies the effect of a severe environment only to find that, in doing so, it has made possible an invasion by a competitor that brings about its own local extinction.

Take, for example, a habitat which is close to the hearts of all students of succession — the sand dune. In West European dunes, the grass *Agropyron junceiforme* is the most conspicuous primary colonist of sands around high-water mark, being tolerant of high salinity (up to 6 per cent according to Salisbury¹) and prolonged immersion. As it grows it creates wind eddies and encourages local sand deposition, and so raises minor dunes around itself. The sand surface rises further and further above the sea, until *Ammophila arenaria*, the marram grass, which has a lower salinity tolerance (about 2 per cent), is able to invade. The higher stature and aggressive, tussocky growth of the marram grass leads rapidly to its dominance. The subsequent decline of *Agropyron* seems simply to reflect its failure in the struggle for survival under competitive conditions.

At this stage many annual plant species invade and much consideration has been given to the factors which limit their establishment^{1–3}. The young dune-front is exposed to adverse conditions but they become less severe as succession proceeds; salt-spray, shifting and unstable sands, low water availability and high diurnal temperature amplitude all may play their part.

Van der Valk⁴, for example, found the seedlings of invading forbs to be more sensitive to sand accretion than to any of the other above factors. In Cape Hatteras, North Carolina, 20–30 cm of winter sand is deposited on the coastal foredunes. But the seedlings can cope with only 5 cm of sand and their seeds with burial by 1–16 cm, so forbs can invade only when the major dominants, such as *Ammophila breviligulata*, have stabilized ridges and produced sheltered locations within the developing undulations and hollows. In the same way, *A. breviligulata*'s European counterpart, *A. arenaria*, produces stabilized dune systems in the European coastal fringe, which support a high diversity of adventive species⁵.

What is very much less clear, because of the difficulty encountered in subjecting any given hypothesis to experimental testing, is why the major dominants like *Agropyron* and *Ammophila* become less

vigorous, flower less frequently and finally die out as the succession proceeds. The obvious answer is that subsequent species have greater competitive ability. But this does not always hold true, for the demise of *Ammophila* in older, stable dunes may not be accompanied by aggressive dominance of the system by robust newcomers. Moribund marram is often surrounded by a short turf vegetation rich in lichens and mosses which do not seem to offer a serious competitive threat; sometimes the declining *Ammophila* populations are accompanied by very little vegetation at all⁶.

Obviously other factors than competition are at work. Salisbury¹ considered increased soil compaction and acidity as possible influences but it was Marshall⁷, whilst working on another sand dune grass, *Corynephorus canescens*, who hit upon a rather surprising answer. He found that burial of *Corynephorus* tufts encouraged the growth of new adventitious roots from tillers of grass and caused its tussocks to spread and proliferate. It grew well under conditions where sand accretion amounted to 10 cm per annum, but became moribund on stable sand and failed to generate new roots. The direct effect of burial was to increase the humidity around the tillers necessary for healthy root growth. Hope-Simpson and Jefferies⁶ confirmed many of the points made by Marshall concerning root growth for *Ammophila*, except for one piece of evidence. When planted in a Bristol garden *Ammophila* grew luxuriantly, both in foliage and root, and flowered freely despite the stability of the soil.

Recently, interest has developed in another dominant plant of dune successions, the sea buckthorn, *Hippophaë rhamnoides*. It is undoubtedly a native plant in Britain⁸, but its frequency in coastal dunes, especially in the west, is probably a consequence of deliberate planting to stabilize sand⁹. It is a shrub which lives for about 40 years¹⁰ and has root nodules in which symbiotic nitrogen fixation occurs¹¹. Its growth can, therefore, have considerable ecological consequences and raise ecosystem nitrogen levels by up to $179 \text{ kg ha}^{-1} \text{ yr}^{-1}$ in sand dunes¹². Stewart and Pearson, who conducted studies on the nitrogen-fixing capacity of *Hippophaë*, noted that nodules on the youngest plants were most active in fixation.

On dunes in the Netherlands, *Hippophaë* enters a degenerate phase in older dunes, in a similar way to *Ammophila*. Oremus and Otten¹³ grew vigorous, young *Hippophaë* in soils from young embryo dunes and from stable sites where the *Hippophaë* was in a post-optimal state of development. More root nodules

- Canal, R., Isern, J. & Labay, J. *Nature* **296**, 225 (1982).
- Bethe, H. A. *Nucl. Phys.* **A324**, 487 (1979).
- Supernova, Proc. NATO Conf. on Supernovae*, Cambridge (eds. Rees, M. J. & Stoneham, R. J.) (Reidel, Dordrecht, in the press).
- Schramm, D. N. & Wilson, J. R. *LLL Preprint* (1982).
- Kolb, E. W., Tubbs, D. & Dicus, D. A. *Astrophys. J. Lett.* (in the press).
- Colgate, S. A. & Petschek, A. G. *Astrophys. J. Lett.* **L115** (1980).
- Lattimer, J. M. & Mazurek, T. J. *Astrophys. J.* **246**, 955 (1981).
- Mazurek, T. J. & Wheeler, J. C. *Fundamentals Cosmic Phys.* **5**, 193 (1980).
- Tinsley, B. M. *Proc. Texas Workshop on Type I Supernovae*, Austin, Texas (1980).
- Weaver, T. A., Axelrod, T. S. & Woosley, S. E. *Proc. Texas Workshop on Type I Supernovae*, Austin, (1980).
- Axelrod, T. S. *Proc. Texas Workshop on Type I Supernovae*, Austin, p. 80 (1980).
- Riegler, G. R. *et al. Astrophys. J. Lett.* **248**, L13 (1981).
- Leventhal, M. & MacCallum, C. J. *Proc. Workshop on the Galactic Center*, Cal Tech (1982).
- Slattery, W. L., Doolen, G. D. & DeWitt, H. E. *Phys. Rev.* **A21**, 2087 (1980).
- Baym, M. G. & Pines, D. *Ann. Phys.* **66**, 816 (1971).

Peter D. Moore is a Senior Lecturer in the Department of Plant Sciences, University of London King's College, 68 Half Moon Lane, London SE24 9JF.

were produced in the embryo dune soil and there was a higher rate of nodule death and necrosis in the mature dune soil. Root growth generally was poorer in the mature soil, roots were darker and root hairs fewer. Gamma irradiation of the soil samples before planting removed these root effects and healthy root systems developed on the plants even when grown in mature dune soils, which suggests that an organism within the soil was responsible for root stunting. A possible culprit was considered — the nematode *Longidorus*. It is absent in the embryo dune soil, but common in that of the mature dune and belongs to a genus known to cause root damage among many plants.

Several possible developments suggest themselves from this work. First, this knowledge of the susceptibility of *Hippophaë* to *Longidorus* infection, if confirmed, could provide a technique for the biological control of *Hippophaë* in those sand dune nature reserves where sea buckthorn has proved a pest by forming such dense and invasive patches that overall species diversity has been

reduced¹⁰. Second, the role of soil fauna in other cases where vigour declines as soils mature, such as *Ammophila*, needs to be investigated. It could account for the healthy growth of marram in the stable gardens of Bristol. Third, the work should serve to underline the potential importance of the soil biota, both faunal and microbial, in influencing competitive interactions and succession in the field. □

1. Salisbury, E.J. *Downs and Dunes: Their Plant Life and its Environment* (Bell, London, 1952).
2. Oosting, H.J. & Billings, D.W. *Ecology* 23, 131 (1942).
3. Chapman, V.J. *Coastal Vegetation* (Pergamon, Oxford, 1964).
4. van der Valk, A.G. *Can. J. Bot.* 52, 1057 (1974).
5. Hope-Simpson, J.F. *J. Ecol.* 53, 447 (1965).
6. Hope-Simpson, J.F. & Jefferies, R.L. *J. Ecol.* 54, 271 (1966).
7. Marshall, J.K. *J. Ecol.* 53, 447 (1965).
8. Godwin, H. *The History of the British Flora* (Cambridge University Press, 1975).
9. Groves, E.W. *Proc. bot. Soc. Br. Isl.* 3, 1 (1958).
10. Ranwell, D.S. (ed.) *The Management of Sea Buckthorn, Hippophaë rhamnoides L. on Selected Sites in Great Britain* (NERC, Norwich, 1972).
11. Bond, G., MacConnell, J.T. & McCullum, A.H. *Ann. Bot.* 20, 510 (1956).
12. Stewart, W.D.P. & Pearson, M.C. *Plant Soil* 26, 348 (1967).
13. Oremus, P.A.I. & Otten, H. *Plant Soil* 63, 317 (1981).

Somatic gene amplification during *Drosophila* oogenesis

from David Ish-Horowicz

THE construction of an egg can make large demands on macromolecular synthesis, especially for organisms in which oogenesis is particularly rapid. Such is the case for *Drosophila melanogaster*, in which oogenesis takes only 2.5 days and as many as two to three eggs an hour can be laid. The intense synthesis of oocyte components is mainly achieved in cells whose DNA content has been amplified up to 1,000-fold by an increase in chromatid number per chromosome (polytenization). Two recent papers^{1,2} show that, in addition to a 45-fold polytenization of the chromosomes in the follicle cells producing the eggshell, there is a further novel mechanism to boost synthesis: a localized differential polytenization of the gene clusters coding for eggshell components.

The protective eggshell, or chorion, contains about 6 major and 14 minor proteins that are synthesized over about 5 hours in a characteristic order that reflects the shell's complex multilayered structure^{3,4}. Two years ago Spradling and Mahowald⁵ showed that the chorion genes are selectively amplified in follicle cell nuclei but not in other tissues. Using cloned cDNA and genomic DNA fragments two clusters of chorion genes were studied: one mapped to the X chromosome within

bands 7F1,2 and includes the genes for proteins s36-1 and s38-1 and two related but uncharacterized genes⁵⁻⁷; the other cluster derives from the third chromosome (between bands 66D11 and -15) and encodes proteins s15-1 and s18-1 and includes two other genes^{7,8}. Each cluster is expressed for only 1–2 hours and appears to be coordinately regulated^{4,6}.

Comparison of egg chamber DNA with that from other male and female tissues showed⁵ that, in addition to the polytenization, the s36/s38 cluster is amplified about 12-fold, and the s15/s18 cluster is amplified about 60-fold. Differential replication is maximal at stages 11–13, the time of chorion protein transcription. At typical rates of transcription, about a 10-fold amplification would be needed to make the observed levels of s36 mRNA in 2 hours⁵, suggesting the extra genes are necessary to make sufficient chorion proteins.

Two possible models of amplification were suggested: additional extrachromosomal chorion genes, or localized differential polytenization. The two recent papers strongly support the latter explanation.

Spradling¹ examined the regions surrounding each cluster and showed that the amplification extends well beyond the chorion genes. Even as far as 40 kilobases (kb) away from the cluster, he finds two- to fourfold amplification. Maximal amplifi-

cation occurs at the chorion genes, with a steady decline in more distant regions, consistent with a replication origin mapping close to the cluster. The steady decrease in amplification suggests the replication terminates randomly rather than at occasional discrete sites. He confirmed the absence of specific termination sites by using Southern transfer hybridization to look for restriction fragments with an altered electrophoretic mobility. No such differences between amplified and non-amplified DNA were found.

Elegant confirmation of the model is presented in the second paper² where Spradling and Mahowald studied a female-sterile mutation *occelliless* (*oc*)⁹ caused by a 10-band chromosomal inversion which breaks the s36/s38 cluster, transposing most of it next to sequences in 8A1,2. The mutation drastically alters the pattern of amplification, as if the inversion has transposed a *cis*-acting replication origin. The inverted chorion genes are still amplified and the replication spreads into the adjacent and previously unamplified 8A1,2 sequences. In contrast, the uninverted 7F1,2 sequences (including at least one chorion gene) are no longer amplified. The replication origin must thus be in the proximal part of the cluster and, indeed, this is the first strong evidence for a discrete eukaryotic chromosomal replication origin. As well as disrupting the pattern of amplification, the *oc* inversion reduces s36 and s38 amplification twofold by inhibition of the last round of somatic amplification. The consequent reduction of s36 and s38 synthesis may contribute to the female sterility (which is due to defective eggshell production¹⁰), but other reasons may include the inability to amplify the untransposed chorion gene.

Is tissue-specific gene amplification during development a common mechanism for selectively enhancing gene expression? This may be the case in *Drosophila* where many, if not most, differentiated tissues

1. Spradling, A.C. *Cell* 27, 193 (1981).
2. Spradling, A.C. & Mahowald, A.P. *Cell* 16, 203 (1979).
3. Petri, W.H., Wyman, A.R. & Kafatos, F.C. *Devl Biol* 49, 185 (1976).
4. Waring, G.L. & Mahowald, A.P. *Cell* 16, 599 (1979).
5. Spradling, A.C. & Mahowald, A.P. *Proc. natn. Acad. Sci. U.S.A.* 77, 1096 (1980).
6. Spradling, A.C. & Mahowald, A.P. *Cell* 16, 589 (1979).
7. Spradling, A.C. *et al.* *Cell* 19, 905 (1980).
8. Griffin-Shea, R. *et al.* *Cell* 19, 915 (1980).
9. Spradling, A.C., Waring, G.L. & Mahowald, A.P. *Cell* 16, 609 (1979).
10. Underwood, E.M. & Mahowald, A.P. *Devl Genet.* 1, 247 (1980).
11. Breuer, M.E. & Pavan, C. *Chromosoma* 7, 371 (1955).
12. Stocker, A.J. & Pavan, C. *Chromosoma* 45, 295 (1974).
13. Postlethwaite, J.H. & Handler, A.M. *J. Insect Physiol.* 25, 455 (1979).
14. Postlethwaite, J.H., Bownes, M. & Jowett, T. *Devl Biol.* 79, 379 (1980).
15. Ganen, A., Kauvar, L. & Lepesant, J.-A. *Proc. natn. Acad. Sci. U.S.A.* 74, 5099 (1977).
16. Audit-Lamour, C. & Busson, D.J. *Insect Physiol.* 27, 829 (1981).
17. Bownes, M. Personal Communication.
18. Petri, W.H. *et al.* *Wilhelm Roux Arch. Entw. Mech. Org.* 186, 351 (1979).
19. Nadel, M.R. & Kafatos, F.C. *Devl Biol.* 75, 26 (1980).
20. Goldsmith, M.R. & Clermont-Rattner, E. *Genetics* 96, 201 (1980).

David Ish-Horowicz is in the Imperial Cancer Research Fund, Mill Hill Laboratories, Burtonhole Lane, London NW7 1AD.

are polytene. A long-standing example comes from the salivary glands of *Rhynchosciara* where DNA amplification of a chromosome band (DNA puffing) precedes its transcription¹¹. In this case, the trigger for DNA replication is the steroid hormone ecdysone¹². Perhaps chorion gene amplification in *Drosophila* is also under hormonal control although there is only suggestive evidence for this. Both ecdysone and juvenile hormone have been implicated in regulating vitellogenesis in *Drosophila*^{13,14}, which occurs at the same stages as the chorion gene replication. *Drosophila* mutants that are temperature sensitive for larval ecdysone production are also temperature sensitive for female fertility and eggshell production^{15,16} but the latter effect appears associated with follicle

cell degeneration¹⁷. As chorion maturation has been demonstrated in cultured ovaries¹⁸, the role of hormones in triggering somatic chorion gene replication is now accessible to *in vitro* study.

It is doubtful that tissue-specific gene amplification is very general. The chorion of the silkworm *Bombyx mori* is also secreted by surrounding follicular cells, but its proteins are made from over 100 linked genes^{19,20} that are not thought to be somatically amplified. Indeed, a host of other intensely transcribed genes are single copy in expressing cells. Presumably, the rarity of somatic gene amplification reflects the ability of a single protein-coding gene to cope with heavy synthetic requirements, except under the most demanding circumstances. □

Some like it hot: temperature-determined sex

from Paul H. Harvey and Montgomery Slatkin

MOST biologists are familiar with chromosomal sex determination. Mammals with both an X and a Y chromosome develop as males while those with two X chromosomes develop as females. In birds the arrangement is reversed and females are the heterogametic sex. Some reptiles, amphibians and fishes also possess sex chromosomes which determine whether the zygote develops as a male or as a female¹.

Genetically controlled determination has generally been assumed to have a remote ancestry in the vertebrate lineage, but recently a quite different sex-determining mechanism has been discovered. Sex in many reptiles is now known to be not genetically controlled but determined by the incubation temperature of the zygote during embryogenesis².

The first case of temperature-determined sex was recognized in 1966 for a lizard³, and others appeared among turtles in the 1970s^{4,5}, but it was not until 1979 that Bull and Vogt⁶ ruled out the possibility that differential mortality dependent upon temperature was the source of the sex-ratio bias. In their study of five species of turtle, they were further able to show that, in four of the species, sex ratios in the wild resulted from the degree of exposure of the nest to sunlight and hence to the incubation temperatures. In the fifth species, sex determination was not temperature dependent. Among reptiles, temperature-dependent sex determination is now known or suspected in alligators, two species of lizards and many turtles² (including sea turtles⁷ — with important implications for

incubation of eggs in the laboratory as a conservation measure for endangered species). Most reptiles have yet to be studied.

Incubation temperatures can have a dramatic effect on sex-ratio bias. In many turtles, low temperatures (20–27°C) produce only males, high temperatures (30–35°C) produce only females and there is a narrow temperature range producing both sexes. Lizards show the opposite pattern with low temperatures causing eggs to develop as females, and high temperatures as males. Snapping turtles (*Chelydra serpentina*) show yet a different pattern of response: females develop at high and low temperatures (20 and 30°C) and males in between. But natural nest sites fluctuate in temperature. In a laboratory experiment subjecting map turtle (*Graptemys*) eggs to daily temperature fluctuations between different all-male and all-female extremes (20–30°C and 23–33°C)⁶, the broods raised between the lower temperatures produced only males, and those between the higher temperatures only females.

In species which have heteromorphic sex chromosomes (those recognizable under a microscope), it seems that the environment does not also determine sex, that is, individuals which deviate from the usual association of XX with one sex and XY with the other are not observed (for a review, see ref. 2). There may, however, be a genetic component to temperature-determined sex. In the Atlantic silverside (*Menidia menidia*)⁸, low temperatures produce more females but eggs from the broods of different females show different temperature-response curves. Similar, though less extreme effects have been noted in turtles⁹.

Little is yet known about how tempera-

ture controls sex determination. The sensitive period lasts some weeks during the middle of development in turtles^{10–12} and occurs about 46 days after hatching in the Atlantic silverside. Temperature may thus be acting during gonadogenesis. H-Y antigen, a cell-surface component suspected of having a primary role in gonadal sex determination in vertebrates (the heterogametic sex is H-Y positive)¹³, has been studied in fourteen turtle species¹⁴. In thirteen species the female was typed as H-Y positive, and in the other species (*Chinemys reevesi*) the male was H-Y positive. Unfortunately, very few individuals of each sex were typed (between one and three) and so the data do not provide firm evidence for a completely sex-differentiated response. Indeed, in two species the male was not typed and a previous study had shown that, in a third species (*Emys orbicularis*), both males and females could be either H-Y positive or H-Y negative¹⁵. This species shows temperature-dependent sex determination, so the results suggest that H-Y antigen does not control sex differentiation in this species. If the findings are correct and eventually hold for gonadal cells (the quoted experiment was performed on blood cells), this constitutes a major departure from the supposed ubiquitous role of H-Y antigen in vertebrate sex determination.

Temperature-dependent sex determination poses both demographic and evolutionary puzzles. What, if anything, prevents large variations in sex ratio in both space and time as developing broods respond to different climatic conditions? Bull and Vogt¹⁶ compared sex-determining temperatures among embryos of map and painted turtles collected from the northern and southern United States and predicted that embryos from southern populations should develop as males at higher temperatures than those of northern populations. Interspecies comparisons provided no support for, and intraspecific comparisons (on two species where this is possible) actually went against, the hypothesis. If this is generally true, then there is no local adjustment of the temperature-dependent mechanism through zygotic means. Populations will then exhibit large geographical differences in sex ratio unless some compensatory mechanism, such as varying the choice of nest sites or emigration, adjusts the adult sex ratio.

Why has temperature-dependent sex determination evolved? If environmental conditions cause a large excess of one sex, then, following Fisher¹⁷, any mechanisms that reduced the excess would be favoured by natural selection. Charnov and Bull¹⁸ have argued that environmental sex determination (of which temperature-determined sex is one case) would evolve when conditions experienced during some stage of development have a strong differential influence on the fitnesses of

Paul H. Harvey is a lecturer in Biology at the University of Sussex, Falmer, Brighton and Montgomery Slatkin is an Associate Professor in Zoology at the University of Washington, Seattle, Washington 98195.

adult males and females. For example, larval mermitid nematodes become female in response to high levels of nutrition (in large hosts), but become male under low levels of nutrition. The explanation is that the large individuals, that can develop in large hosts, can produce more offspring as females than as males. At present, it is not clear whether such effects exist in reptiles but a new study of alligators (see this issue of *Nature*, p.850) suggests that they may. Incubation temperature indirectly affects the animals' growth rate by influencing the amount of yolk remaining at hatching. Large size early in life is apparently more beneficial to females than to males.

Temperature-determined sex may indeed be the ancestral condition for reptiles. It is known in three of the four living orders: turtles, snakes and lizards, and crocodilians (the tuatara, the surviving representative of the Rhynchocephalia, has not been studied). Morphologically distinguishable sex chromosomes are found in only two orders (the turtles and the snakes and lizards), but it is uncertain

whether genotypic sex determination originated once or more than once². If temperature-determined sex is ancestral for reptiles, and if the birds and mammals arose from such a reptilian stock, then we are left with a radically different view of the evolution of sex determination in the vertebrates. □

1. Ohno, S. *Sex Chromosomes and Sex-Linked Genes* (Springer, Berlin, 1967).
2. Bull, J.J. *Q. Rev. Biol.* **55**, 3 (1980).
3. Charnier, M. *Soc. Biol. Quest. Af.* **160**, 620 (1966).
4. Pieau, C. *Bull. Soc. zool. Fr.* **100**, 67 (1975).
5. Yntema, C.L. *J. Morph.* **150**, 453 (1976).
6. Bull, J.J. & Vogt, R.C. *Science* **206**, 1186 (1979).
7. Yntema, C.L. & Mrosovsky, N. *Herpetologica* **36**, 33 (1980).
8. Conover, D.O. & Kynard, B.E. *Science* **213**, 577 (1981).
9. Bull, J.J., Vogt, R.C. & Bulmer, M.G. *Evolution* (in the press).
10. Yntema, C.L. *J. Morph.* **159**, 17 (1979).
11. Bull, J.J. & Vogt, R.C. *J. exp. Zool.* **218**, 435 (1981).
12. Pieau, C. & Dorizzi, M. *J. Morph.* **170**, 373 (1981).
13. Wachtel, S.S., Koo, G.C. & Boyse, E.A. *Nature* **254**, 270 (1975).
14. Engel, W., Klemme, B. & Schmid, M. *Differentiation* **20**, 152 (1981).
15. Pieau, C., Dorizzi, M. & Zaborski, P. *Archs Anat. microsc. Morph. exp.* **68**, 22 (1979).
16. Bull, J.J. & Vogt, R.C. (cited in ref.2); also Bull, J.J., Vogt, R.C. & McCoy, C.J. *Evolution* (in the press).
17. Fisher, R.A. *The Genetical Theory of Natural Selection* (Oxford University Press, 1930).
18. Charnov, E.L. & Bull, J.J. *Nature* **286**, 828 (1977).

An Arctic Ocean ice sheet in the Pleistocene?

from L.D. Keigwin

AN intriguing and controversial suggestion stemming from recent palaeoclimatic studies is that the Arctic Ocean was covered by a massive floating ice shelf during the last glacial maximum. The idea has been put forward from evidence of $\delta^{18}\text{O}$ fluctuations in deep-sea cores and in ice cores on Greenland¹ and has been proposed as a solution to problems in Arctic glacial dynamics and chronology².

A recent paper by Williams, Moore and Fillon³ now puts the floating Arctic Ocean ice shelf hypothesis in a time-stratigraphical framework for the entire Pleistocene, and for the last 150,000 years in particular, by comparing inferred sea levels from the $\delta^{18}\text{O}$ records of the deep sea with the record of sea level given by raised coral terraces. They found significant differences in the apparent sea level calculated from the two data sources which they attributed to the presence of a floating ice shelf that would produce the change in seawater isotopic composition without significantly altering sea levels. The authors conclude that ice shelves may have originated as important features earlier in the Quaternary.

Oceanic $\delta^{18}\text{O}$ records have been observed to change from high-frequency, low-amplitude to low-frequency, high-amplitude variability about 900,000 years before present. Williams *et al.* argue that at that time the Earth's orbital parameters

changed and allowed the first combined accumulation of massive ice shelves and sheets in the Arctic.

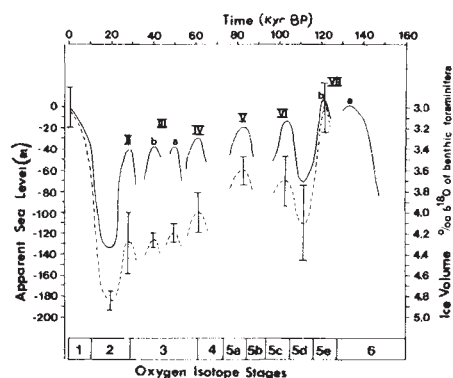
For the Quaternary it is widely held that $\delta^{18}\text{O}$ changes of 0.10‰ in deep-sea benthic foraminifera correspond to sea-level changes of 10 m (based on a mass balance assuming average $\delta^{18}\text{O}$ values of glacial ice⁴). A similar relationship has been derived from $\delta^{18}\text{O}$ variations in corals from uplifted terraces⁵. The only known way rapidly and greatly to enrich the ocean in ^{18}O is by removing isotopically lighter ^{16}O from the ocean through evaporation and precipitation. A fall in sea level will result if the precipitation is stored on continents as ice, but not if it is stored as a floating ice shelf. The $\delta^{18}\text{O}$ of deep-sea benthic foraminiferal calcite depends on the temperature and isotopic composition of the seawater and so-called 'vital effects', but most of the 1.7‰ glacial-interglacial $\delta^{18}\text{O}$ change in the ocean can be attributed to seawater compositional changes caused by increased ice volume. Vital effects can be reduced by analysing single species of benthic foraminifera and the effect of changing deep-sea temperatures on $\delta^{18}\text{O}$ records is generally discounted because the total amount of isotopic change would

represent a cooling of most of the deep sea below freezing.

Sea-level records based on raised coral terraces are not continuous, but can be dated. Constant rates of uplift have been assumed and, where reasonable, a composite sea-level record has been derived. Williams *et al.* compared such a composite terrace record with a composite deep-sea record for the late Quaternary³ to arrive at their conclusions. Other sea-level indicators, however, do not always agree with the raised coral terrace record.

Hollin⁶ has provided evidence for a dramatic rise in sea level about 95,000 years ago which may be due to an Antarctic glacial surge. If such events occur in less than the 1,000 year mixing time of the ocean they may not be resolved in the deep-sea stable isotopic record. The greatest apparent sea-level differences between the generalized records of Williams *et al.*³ are between about 20,000 and 100,000 years ago where they differ by 30–80 m. To produce a 50 m sea-level difference the area of the Arctic Ocean would have to be covered by a floating ice shelf 1,300 m thick, close to the mean depth of that ocean. Even if adjacent seas and part of the North Atlantic are included, hundreds of metres of ice shelf are still required.

Alternatively, the 50 m sea-level difference (equal to a difference of $\sim 0.50\text{‰}$) could be accounted for by a deep-ocean cooling of 2°C. Williams *et al.* have not given this possibility emphasis because of ambiguities in inferring bottom-water mass history from benthic foraminiferal distributions. Studies of cores under Antarctic bottom water show faunal changes that do not follow a simple glacial-interglacial pattern^{7,8}, but studies of cores beneath deep waters (Circumpolar and



Comparison of apparent sea-level records from raised coral terraces (solid line) and benthic foraminiferal oxygen isotope records (dashed line). The isotopic record is a composite of results from five Atlantic, Indian and Pacific Ocean deep-sea cores and error bars represent one standard deviation about the mean at various times. Vertical scales are adjusted so that a 1.1‰ $\delta^{18}\text{O}$ change equals a 10 m sea-level change. Sea-level differences between the two records of as much as 80 m may result from ice shelves which could change the isotopic composition of the ocean without significantly lowering sea level, and from cooler deep-sea temperatures during glacial intervals.

L.D. Keigwin is at the Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543.

North Atlantic deep water) do show glacial-interglacial faunal oscillations, suggesting water mass changes^{9,10}. The isotopic data examined by Williams *et al.* are based on sediment from cores of relatively shallow water depth (< 3,300 m) within deep-water masses. Deeper cores generally have lower sedimentation rates because calcite dissolution increases with depth; thus they have lower resolution. As such, most cores from which isotopic data have been reported actually could have 2°C glacial-interglacial changes in deep-water temperature.

Deep-sea cooling could be eliminated from consideration by studying stable isotopes in cores where deep-sea water is less than 2°C from freezing. Water mass changes could also significantly affect the interpretation of benthic foraminiferal isotopic records as ice volume-sea level indicators because the range of the principal modern deep-water masses is about 0.6‰. For example, a seawater compositional change of this size could result from replacing North Atlantic deep water by Antarctic bottom water at the transition from an interglacial to a glacial interval. Recent stable isotopic and geochemical studies of cores within the Atlantic suggest reduction or elimination of North Atlantic deep water during glacial intervals¹¹⁻¹⁴.

Regardless of the uncertainties in interpreting foraminiferal $\delta^{18}\text{O}$ data, some glaciologists have hypothesized that during glacial extremes there were massive ice shelves associated with an Arctic ice sheet^{2,15,16}. This controversial hypothesis is based on analogy with the Antarctic ice sheet, a large part of which is currently grounded below sea level. Central to the idea² is the role of ice shelves in explaining the location, mode and sequence of formation of most Northern Hemisphere ice domes. Briefly, Hughes *et al.*² propose that massive ice shelves form when ice streams coalesce and the shelves buttress and allow the growth of massive domes. The shelves must form before the domes, so the largest ice volume growth in the late Quaternary would require ice shelves.

The work of Williams *et al.* has focused attention on the debate about the extent of floating Arctic ice during glacial intervals. Although evidence from Arctic Ocean cores indicates perennial sea-ice cover began 700,000 to 900,000 years ago¹⁷, it is

difficult enough to prove the existence of massive ice shelves during the last glaciation, not to mention earlier in the Quaternary. Further studies of deep-sea cores, including cores from the Arctic may help resolve the problem. □

The key to the schistosome's success

from A.J.G. Simpson and D. Cioli

SCHISTOSOMIASIS or bilharzia is a disease affecting some 200 million people in the developing countries. The causal agent is a parasitic worm, the schistosome, which lives in the blood vessels of the intestine or urinary tract; man becomes infected by the cercarial, or free-living aquatic, stage which penetrates the skin directly. The key to understanding how host immunity is effected and how the worm evades this response is to learn more about the schistosome's double outer membrane (DOM). This unique surface not only mediates crucial metabolic functions but also counteracts the immune system of the host. Progress in this field was reported at a recent meeting in Geneva*.

The techniques of subcellular fractionation are providing an invaluable approach to the characterization of the DOM. It can be selectively released from the parasite simply by brief incubation in buffered saline at 37°C (Evans, NIMR, London), by using freeze-fracture techniques (Wilson, University of York), or by utilizing the negative surface charge of the parasite as the basis of membrane purification (Cesari, Instituto Venezolano de Investigaciones, Caracas). An alkaline phosphatase of broad specificity has been widely used as a marker for the DOM and enrichment of specific activity for this enzyme has been reported, by all investigators, in fractions shown to contain DOM by morphological observation or in those enriched in externally applied ligands.

Wilson and Podesta (University of Western Ontario) have further dissected the DOM by preparing separate fractions enriched for the inner and outer bilayers. Podesta utilized the sequential exposure of the parasite to 0.1 per cent digitonin solutions to obtain first the outer and then the inner bilayer from intact adult worms, a technique previously used in the preparation of mitochondrial membranes in other systems. The outer bilayer fractions had very low enzyme activities, suggesting that membrane-bound enzymes are associated predominantly with the inner layer of the DOM.

The surface of the schistosome represents the primary target for immune attack. As the schistosomulum (the immature worm) is the most susceptible stage, attempts to determine target antigens have largely been conducted by surface labelling this stage. Several participants reported results carried out by a variety of techniques. As many as ten or more proteins of molecular weight 18,000–200,000 may be labelled on schistosomula. There is, however, variation in the polypeptides labelled by the various procedures, and Cioli (Consiglio Nazionale della Ricerche, Rome) has shown that the most selective and gentle methods identify an 18,000 molecular weight polypeptide. Not all labelled polypeptides can be immunoprecipitated with sera from patients or infected animals and the reported molecular weights of major precipitated antigens vary considerably.

Taylor (University of Cambridge) used iodosulphanilic acid to label the schistosomulum and immunoprecipitated polypeptides of molecular weights 185,000, 105,000, 68,000 and 24,000 with sera from infected humans. Dissous (Institut Pasteur, Lille) immunoprecipitated polypeptides of 40,000, 37,000 and 32,000 with immune rat sera following lactoperoxidase-catalysed iodination of the schistosomular surface. Both of these investigators reported the production of monoclonal antibodies which recognized defined surface antigens (Taylor — molecular weight 24,000, Dissous — 37,000) and detected some *in vitro* activity with their monoclonal reagents. Dissous further reported the passive transfer of immunity to rats using her monoclonal antibody.

How is the DOM of the developing parasite involved in immune evasion? Susceptibility to immunity is age dependent and as the parasite matures it quickly loses the ability to bind anti-parasite antibody, a change which correlates with the detection of host deter-

1. Broecker, W.S. *Science* **188**, 1116 (1975).
2. Hughes, T., Denton, G.H. & Grosswald, M.G. *Nature* **266**, 596 (1977).
3. Williams, D.F., Moore, W.S. & Fillon, R.H. *Earth planet. Sci. Lett.* **56**, 157 (1981).
4. Shackleton, N.J. *Phil. Trans. R. Soc. B280*, 169 (1977).
5. Fairbanks, R.G. & Matthews, R.K. *Quat. Res.* **10**, 181 (1978).
6. Hollin, J.T. *Nature* **283**, 629 (1980).
7. Lohmann, G.P. *Oceanus* **21**, 58 (1978).
8. Corliss, B.H. *Quat. Res.* **12**, 271 (1979).
9. Streeter, S.S. & Shackleton, N.J. *Science* **203**, 168 (1979).
10. Corliss, B.H. *Geol. Soc. Am. Abstr. with Programs* **12**, 406 (1980).
11. Duplessy, J.-C., Mayes, J. & Pujol, C. *Nature* **286**, 479 (1980).
12. Curry, W.B. & Lohmann, G.P. *Geol. Soc. Am. Abstr. with Programs* **13**, 435 (1981).
13. Curry, W.B. & Lohmann, G.P. *EOS* **62**, 904 (1981).
14. Boyle, E.A. & Keigwin, L.D. *EOS* **62**, 904 (1981).
15. Mercer, J.H. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **8**, 19 (1970).
16. Denton, G.H. & Hughes, T.J. *The Last Great Ice Sheets*, 484 (Wiley, New York, 1981).
17. Herman, Y. & Hopkins, D.M. *Science* **209**, 557 (1980).

*This meeting was organized by the Scientific Working Group on Schistosomiasis of the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases. A fuller report may be obtained from the Secretary, SWG on Schistosomiasis, World Health Organization, 1211 Geneva 27.

A.J.G. Simpson is at the Laboratory of Parasitic Diseases, Bldg 5, NIH, Bethesda, Maryland 20205, and D. Cioli is at the Laboratorio de Biologie Cellulare, Consiglio Nazionale della Ricerche, Via G. Romagnosi 18/A, 00196 Rome.

minants on the parasite surface. Caulfield (Harvard Medical School) reported, however, that parasite antigens are spontaneously lost from young schistosomula *in vitro*, even in the absence of serum or other macromolecules. Thus, the loss of antigenicity may result from the active shedding of parasite surface components.

Membrane turnover and shedding (particularly in response to ligand and cell binding) may be generally involved in the protection of the parasites. Brink (Harvard Medical School) demonstrated the shedding of antibody-antigen complexes from *in vitro* cultured schistosomula, and Caulfield reported that following a unique membrane fusion between the schistosomulum and neutrophil membranes, whole areas of membrane may be lost to the cell interior, but the parasite survives and eventually the neutrophils are released. Kemp (Texas A & M University) has also detected immunoglobulins bound to the adult surface via their Fc portion, and if anti-Fab antibodies are bound to the surface immunoglobulins, the complexes are actively shed within 15 minutes. Freeze-fracture (Torpier, Institute Pasteur, Lille) shows that this is accompanied by the cross-linking of membrane components (visualized as intramembranous particles) into tetrads and their subsequent disappearance from the DOM.

Alterations in the chemical composition of the double outer membrane may result in its being refractory to the effects of cytotoxic mechanisms. When five-day schistosomula are recovered from mice

(McLaren, NIMR, London) they are coated with red blood cell antigens. If the parasite is exposed to eosinophils in the presence of anti-mouse red blood cell antibody, the cells bind to the surface and degranulate, but fail to kill the parasite although similar antibody-dependent cytotoxic mechanisms are potent killers of fresh schistosomula, and also of parasites three weeks old or more. Thus, even if the surface components of five-day old worms are recognized by antibody and cells, there is still an 'intrinsic' protection of the surface membrane of the parasite at this stage.

The presence of host antigen determinants on the parasite surface may be the result of an active evasion mechanism by the schistosome, whereby it synthesizes molecules that mimic those of the host and serve as a disguise. Among the molecules recognized on parasites recovered from mice are the products of the major histocompatibility complex. Simpson (NIH, Bethesda) reported that monoclonal reagents indicate that an intact H-2 antigen is present on the schistosome surface. Although analysis of the parasite genome demonstrated some DNA sequence homology between *S. mansoni* and its hosts, genes coding for either human or mouse histocompatibility antigens were not present, so these molecules must be acquired from the host.

In vitro experiments (Rumjanek, NIMR, London) demonstrated that lipid composition of schistosomula was profoundly affected by the presence of serum which is able to stimulate intense

exchange of lipids between the organism and the medium. Fetal calf serum produces a net depletion of lipids, whereas human serum produces a substantial uptake of cholesterol and triglycerides. This serum-induced lipid modulation correlates with the degree of protection observed in *in vitro* cytotoxicity assays and furthermore, no protection is afforded by incubation of schistosomula in delipidated serum. This strongly implicates lipids as being responsible for the acquisition of resistance by the parasite to the immune effector mechanisms. The protective effect of serum may be associated (Torpier) with low-density lipoproteins, which bind to the parasite surface, and this may be the basis of the lipid changes observed by Rumjanek.

It is clear that our knowledge of the physiological functions of the schistosome DOM is still very limited, but in view of the importance that understanding of the membrane may have for the development of immuno-prophylactic and chemotherapeutic agents, it was felt that increased research efforts should be made to define the molecular structure of specific membrane proteins. The use of monoclonal antibodies and recombinant DNA technology to clone and sequence the parasite genes coding for these components is most likely to provide the way forward. The expression in microorganisms of genes coding for surface antigens may be an important step towards the production of significant amounts of parasite antigens for vaccination. □

100 years ago

No group of animals is more characteristic of the peculiar fauna of Australia than the great family of Honey-Eaters (Meliphagidae), of which upwards of sixty species, belonging to many different genera, are distributed throughout the length and breadth of that Continent. But although the Honey-eaters are so common in Australia, and there is an extensive importation of living birds from Sydney and other Australian ports every year into this country, very few of the Meliphagidae have yet reached Europe alive. The fact is, that the organisation of the Honey-eaters, being adapted for an active and wandering life, in perpetual search of the nectar of the flowering-trees which their pencilled tongue so admirably fits them to collect, does not render them very suitable subjects

for captivity, and it is only recently that means have been found to preserve these birds alive and in good health in cages. It has thus happened that almost the only one of the vast tribe of Australian Honey-eaters that has been exhibited in the Zoological Society's aviaries is the present species, which we now figure (Fig. 18) from four examples lately received from New South Wales.

In his great work on the "Birds of Australia", Mr. Gould tells us that the Warty-faced Honey-eater is not only one of the handsomest of its tribe, but also one of the most beautiful birds inhabiting Australia, the strongly contrasted tints of its black and yellow plumage rendering it a most conspicuous and pleasing object, particularly during flight.

Although very generally distributed, its presence appears to be dependent upon the state of the *Eucalypti*, upon the blossoms of which it mainly depends for subsistence; it is consequently only to be found in any particular locality during the season that those trees are in blossom. It generally resorts to the loftiest and most fully-flowered tree, where it frequently reigns supreme, buffeting and driving every other bird away from its immediate neighbourhood.

The Lobed Muskduck (*Biziura lobata*). — Waterfowl have always formed a favourite portion of the Zoological Society's living collection, and it is with great satisfaction that every new addition to the already long list of "acclimatisable" Anatidae is announced in their journals. The species which is now portrayed (Fig. 19) is certainly one of the most remarkable that they have yet procured, and although perhaps not likely to be "acclimatised" at present, is well worth examination as being remarkable even amongst Australian animals, for several very abnormal features in its structure.

The Musk-duck has a lengthened, stiff, and leather-like appendage hanging from the under surface of the bill, and is the only member of the



family which possesses this singular structure. its lengthened tail, composed of twenty-four narrowed and stiffened feathers, is, no doubt, most serviceable to it in swimming and diving. The female does not carry the chin-lobe, and is very much smaller than the male bird.

The Musk-duck is widely distributed on the Australian Continent, and also inhabits Tasmania. It frequents the bays and inlets of the sea, the upper parts of rivers, lakes, and secluded pools. "More than a pair are rarely seen at one time; often a solitary individual takes up its abode in some favourite pool, where it lives a life of complete seclusion, depending for its food and for its preservation from danger upon its powers of diving rather than upon those of flying. It is very difficult to shoot, as it dives instantly a gun is fired, so that the shot has hardly time to reach it. From Nature 25, 608, 27 April, 1882.



REVIEW ARTICLE

Actin-binding proteins—regulators of cell architecture and motility

Alan Weeds

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

Numerous actin-binding proteins from a variety of cell types have been described. Here I attempt to correlate the properties and functions of some of these. Three major classes have been identified: (1) cross-linking proteins which form filament bundles or isotropic gels; (2) proteins which cap filament ends and nucleate the polymerization of G-actin (many of these also sever actin filaments); (3) proteins which bind to G-actin and stabilize the monomer pool. Some of the proteins described here combine the properties of more than one class and the activities of many of them are regulated by changes in Ca^{2+} ion concentration.

ACTIN is a major constituent of many types of eukaryotic cells and is involved in a wide variety of motile processes including cell locomotion, cytoplasmic streaming and transport, secretion, phagocytosis and cytokinesis. 'Stress fibres', consisting of bundles of microfilaments¹ are clearly seen in spreading or moving cells and these contain myosin, filamin, tropomyosin and α -actinin in addition to actin²⁻⁵. Actin is also concentrated as a dense filament meshwork in the cortical cytoplasm^{6,7}, where finger-like projections (filopodia) and veil-like sheets of membrane (lamellipodia or ruffles) display the dynamic activity of the cell surface during movement. This cortical cytoplasm undergoes dramatic changes in viscosity and stiffness which reflect the formation of actin filaments and their assembly into more highly organized meshwork structures. An understanding of cell motility therefore requires detailed knowledge of (1) the structural components involved with actin in the formation and rearrangement of this cortical network; (2) its biochemical composition; and (3) the mechanisms by which the interaction of these individual proteins are regulated.

Over the past five years much effort has been directed towards these goals⁸⁻¹¹. Many new actin-binding proteins have been identified, initially in crude cell-free extracts, but many of these have now been purified and partially characterized. Particular progress has been made with cytoplasmic gels isolated from macrophages and amoeboid cells and with the more highly organized structure of the intestinal microvillus.

Certain unifying principles are now emerging which may ultimately allow assignment of many of these apparently different proteins to a small number of functional groups. Just as the properties of myosins have been refined by natural selection to suit the particular requirements of a given muscle, so too proteins that form gel networks of actin or regulate the dissolution of these gels may have evolved specialized properties in their individual situations yet fall within a similar functional classification. The common theme is the role of actin filaments in maintaining cell architecture and generating movement. Actin-binding proteins provide the necessary variations for the part actin has to play in such widely differing processes as egg fertilization, blood clotting, intestinal absorption and tumour invasion.

Before reviewing these actin-binding proteins, it is important to note that while the roles of actin, myosin and associated proteins in striated muscle contraction may serve as models for certain types of cytoplasmic movement, there are basic differences between these two systems which demand distinctive properties for proteins common to both. The organization of filaments in vertebrate striated muscle is highly uniform and rigidly maintained to carry out its contractile functions, but in

non-muscle systems, filamentous structures are more dynamic, responding both temporally and spatially to the cell requirements. Ca^{2+} ions serve not only to promote acto-myosin interaction, but dissolve gels and in many cases sever actin filaments. For example, recent studies have shown that α -actinin from non-muscle cells cross-links actin filaments only in the absence of Ca^{2+} ions, in contrast to α -actinin in striated and smooth muscles, which shows no comparable Ca^{2+} -sensitivity. (Indeed if it were Ca^{2+} -sensitive, disintegration of the Z-disk structure might occur during muscle activation!)

General classification of actin-binding proteins

Actin is evolutionarily highly conserved in amino acid sequence¹² which probably reflects the strong selective pressure to preserve its three-dimensional structure and thus maintain essential interaction sites for many associated proteins. Central to the role of actin is its self-association from a monomeric form (G-actin, molecular weight (M_r) 42,000) to give polymers several micrometres in length, containing about a thousand monomers. This polymerization process has been studied in detail, particularly by Oosawa and Asakura¹³ and by Wegner¹⁴ and others but is still not fully understood. The rate of filament assembly is controlled by the formation of nuclei and their production is favoured in the presence of Mg^{2+} ions. Thus, at physiological KCl concentration and with MgCl_2 in excess of 1 mM, actin exists predominantly in the polymerized form, in equilibrium with a very small ($<1 \mu\text{M}$) concentration of free monomer (the so-called critical concentration). Monomers can add to both ends of growing filaments: Pollard and Mooseker¹⁵ estimated the rates of assembly and disassembly at each end and have shown that there is preferred assembly at the 'barbed end' while preferential disassembly occurs at the 'pointed end'. (Based on the structure of actin decorated with myosin subfragment-1, these ends have been termed 'barbed' and 'pointed' by analogy with the flights on an arrow.) A consequence of any difference in assembly rates at the two ends is that actin monomers can cycle through the filaments from barbed to pointed end; this has been termed 'treadmilling'¹⁶. ATP hydrolysis is required to maintain this steady-state process. Although treadmilling of tubulin in microtubules has been clearly shown to occur¹⁶, actin monomer exchange *in vitro* appears to be limited in both extent and duration at actin concentrations $>10 \mu\text{M}$ ¹⁷ and steady-state treadmilling has not been demonstrated to date. One consequence of a difference in assembly rates at the two ends is that filament stability is altered if one end is blocked; for example, capping the preferred assembly

end will cause monomer dissociation from the other end to establish a new monomer-polymer equilibrium¹⁸. I shall demonstrate below how capping proteins may affect both filament stability and polarity of assembly.

In many cell extracts, the concentration of monomeric actin is much higher than that expected on the basis of the critical concentration, due to the presence of profilins¹⁹⁻²¹, proteins which bind specifically to G-actin and prevent polymerization. Profilins are representative of a class of proteins that stabilizes the actin monomer pool, although actin depolymerizing factors and even proteins such as DNase I may perform similar functions. (G-actin forms a 1:1 complex with DNase I and inhibits its enzymatic activity²².)

Once filaments have been formed, other proteins are needed to stabilize them in three-dimensional gels or oriented bundles. Several classes of cross-linking proteins have been described. Some of these 'gelation factors' provide cross-links insensitive to changes in Ca^{2+} ion concentration, while others operate only at Ca^{2+} concentrations $<10^{-7}$ M and the links are broken at $>10^{-6}$ M. Proteins that form actin bundles may also exist in either calcium-sensitive or calcium-insensitive forms.

Thus, these proteins act to form three-dimensional isotropic networks or tight bundles of actin filaments which themselves may further associate to produce more organized structures or to interact with membranes or cell organelles. As actin filaments are often found closely apposed to membranes, it is reasonable to suggest that specific linker proteins exist either as integral membrane components or to attach microfilaments to trans-membrane elements. These have proved very difficult to identify.

Actin-containing gels can undergo contraction in conjunction with cytoplasmic myosins and suitable regulatory proteins (kinases and calmodulin). Another effect observed in the presence of Ca^{2+} ions is that of 'solation', where the gels are dispersed. Gel-sol transformations occur in some cases by severing of the actin filaments, and involve a new class of

actin-binding proteins which I term 'gelsolins'. In many cases, the severing protein remains bound to one end of the actin filament, preventing re-annealing, hence these proteins behave additionally as 'capping factors'. In this respect they mimic the effects of cytochalasin drugs, which inhibit elongation of actin filaments by binding to their barbed ends²³⁻²⁶. One possible outcome of such capping is an increased monomer concentration as a result of G-actin dissociation from the free pointed ends. In the presence of profilins, complete filament depolymerization may occur. These proteins have a second and rather different property. When added to G-actin in assembly conditions, they promote the formation of actin nuclei by binding to more than one monomer and thereby facilitate polymerization. Indeed both accelerated nuclei formation and decreased rates of elongation occur simultaneously with macrophage gelsolin *in vitro* (see below). Thus, depending on the conditions, these proteins may be involved in generating new filaments or in severing and disassembling pre-existing ones.

In the following sections I describe some of the proteins belonging to each of these classes, with particular emphasis on mammalian cells.

Actin cross-linking proteins—gelation and bundling

As already stated, there are several proteins which cross-link actin filaments to form either bundles or isotropic gels. In some cases, gel formation is favoured by a low molar ratio of a particular cross-linking protein to actin, while, with increasing amounts of the same cross-linker, parallel alignment of filaments occurs to form bundles. It is not possible, therefore, to designate every cross-linking protein within either the gelation or bundling subclass, but the large flexible proteins like filamin occur predominantly in gels, whereas much smaller proteins like fascin or fimbrin exist in organized actin bundles and are probably not sufficiently flexible to accommodate isotropic gel formation.

Table 1 Classification of actin-binding proteins

	Source	$M_r \times 10^{-3}$	No. of subunits	Ca^{2+} -sensitivity	Ref.
Gelation proteins					
Filamin	Muscle	250	2	—	27
Filamin	Macrophage	270	2	—	31
Spectrin	Red blood cell	240 (α) 220 (β)	4	—	37
—	Sea urchin egg	220	—	—	60
—	<i>Dictyostelium</i>	120	—	—	39
Gelation and bundling proteins					
α -Actinin	Muscle	105	2	—	43
α -Actinin	HeLa cells	100	2	+	49
Actinogelin	Ehrlich ascites	115	—	+	51
—	<i>Dictyostelium</i>	95	—	+	52, 53
—	<i>Acanthamoeba</i>	85	—	+	54
Vinculin	Muscle	130	1	—	45
Vinculin	HeLa cells	130	1	+	50
Bundling proteins					
Fascin	Sea urchin egg	58	—	—	60
Fimbrin	Microvillus	68	1	—	62
Severing and capping proteins					
β -Actinin	Muscle	34 + 37	1 each	—	64
Gelsolin	Macrophage	90	1	+	66
Gelsolin	Platelets	90	1	+	69
Gelsolin	Plasma	90	1	+	73, 75
Villin*	Microvillus	95	1	+	89, 91
Fragmin	<i>Physarum</i>	42	1	+	78, 79
—	<i>Dictyostelium</i>	40	1	+	81
Capping protein	<i>Acanthamoeba</i>	28 + 31	1 each	—	82
Depolymerizing protein	Brain	19	—	—	83
G-actin stabilizing proteins					
Profilin	Lymphocytes	16	1	—	19
Profilin	<i>Acanthamoeba</i>	14	1	—	21

This table includes the main proteins described in the text and omits a large number of well known actin-binding proteins found in muscle and elsewhere. The molecular weights are from polyacrylamide gel electrophoresis and as buffer conditions vary between different laboratories, the values cannot be strictly compared. Comparisons should not assume accuracy better than 5%. Subunit determined from M_r of native protein.

* Also a bundling protein.

Filamins. A high molecular weight actin-binding protein, purified from smooth muscle, has been termed filamin²⁷. A structurally similar protein was identified in lung macrophages²⁸ and subsequently in neutrophils, platelets and other cell types. The general morphology of macrophage actin-binding protein is indistinguishable from filamin²⁹, both appearing as two flexible strands joined at one end, but they differ in their potencies for cross-linking actin³⁰. The overall length of the dimer of $M_r \sim 500,000$ is ~ 160 nm. Here I will use the term 'filamin' for this general class of large flexible actin cross-linking proteins.

Hydrodynamic measurements indicated a native M_r of 540,000 and macrophage filamin bound to actin filaments at 40 nm intervals, giving a stoichiometry of 1 filamin per 15 actin monomers at high molar ratios³¹. Bundles of actin filaments were seen in these conditions *in vitro*, but at lower ratios of filamin to actin, isotropic gels were produced. The concentrations of myosin required to produce contraction of actin gels were lower in the presence of macrophage filamin than in its absence³². When actin was polymerized in the presence of macrophage filamin, branched filaments were produced having an average length less than that seen in the absence of filamin³³. Filamin also accelerates the onset of polymerization and produces a perpendicular branching of actin filaments, thereby promoting the formation of an isotropic gel. It is located in the pseudopods of polymorphonuclear leukocytes during locomotion and phagocytosis³⁴, in stress fibres and membrane ruffles of dividing chick embryo cells during interphase, and appears to be concentrated in the cleavage furrow during cytokinesis³⁵.

The morphological equivalent of filamin in the red blood cell is spectrin. Although spectrin appears as a loose, flexible and elongated dimeric molecule, in contrast to filamin, spectrin dimers are joined at both ends and self-associate end-to-end to form tetramers in 0.1 M KCl. Spectrin tetramers of ~ 200 nm in length cross-link small protofilaments of actin (10–15 subunits long on average) to form the scaffold of proteins that maintains the shape of the erythrocyte. These complexes of spectrin and actin are stabilized by a globular protein, band 4.1 in the Steck classification³⁶. This cytoskeletal framework is attached to the membrane via ankyrin, which links the spectrin tetramers to the main transmembrane protein, band 3 (see refs 37, 38 for reviews). Spectrin-like proteins may not be restricted to red blood cells, but may serve similar functions elsewhere.

Recently, a smaller protein that gels F-actin has been isolated from *Dictyostelium discoideum*^{39,40}. Like filamin, it promotes side-to-side and end-to-side interactions between actin filaments and increases the initial rate of actin polymerization. It also seems to be localized in the cell cortex and surface ruffles. Its native molecular weight and actin-binding stoichiometry have not been characterized.

α -Actinins. While filamin and spectrin are representatives of a class of large flexible actin cross-linking proteins, a distinct class of rod-like actin-binding proteins has been identified in muscle by Ebashi and co-workers⁴¹. The archetype, α -actinin, is located in the Z-disks of muscle and seems to be involved in cross-linking actin filaments⁴². The native protein is a dimer ($2 \times M_r$, 105,000) and the molecule is highly asymmetric, with a length of 400 Å (ref. 43). α -Actinin forms gels with F-actin *in vitro*, although its binding to actin is inhibited by tropomyosin⁴⁴.

In non-muscle cells, antibodies to muscle α -actinin stain stress fibres and regions close to the membrane^{44,45}. These observations suggested that α -actinin is involved in binding microfilaments to membrane surfaces. However, although immunocytochemistry shows that α -actinin is concentrated close to focal adhesion plaques in cultured cells⁴⁵, it does not form a direct link with membranes⁴⁶. Furthermore, α -actinin can be selectively dissociated from plasma membranes with little parallel release of actin⁴⁷. α -Actinin therefore seems to act as a cross-linker and spacer between actin filaments⁴⁸.

An important recent discovery is the Ca^{2+} -sensitivity of α -actinin function. Both striated and smooth muscle α -actinins cross-link actin independently of Ca^{2+} concentration, but the

α -actinin of HeLa cells forms actin gels only at Ca^{2+} ion concentrations $< 10^{-7}$ M (refs 49, 50). This calcium-sensitive cross-linking is similar to that of actinogelin from Ehrlich ascites cells⁵¹. Rod-shaped, Ca^{2+} -sensitive actin cross-linking proteins have also been identified in *Dictyostelium*^{52,53}, *Acanthamoeba*⁵⁴ and blood platelets. These proteins have not been completely characterized and although they may not be related by amino acid sequence or immunological criteria, their general morphology and actin cross-linking properties suggest that they may be classified within the same group. A number of smaller gelation factors have been described in amoeboid cells such as *Dictyostelium* and *Acanthamoeba*, based on viscometric or turbidimetric measurements, but these have yet to be purified.

One further interesting property of α -actinin is that it appears to dissociate the profilin-actin complex *in vitro* and promote polymerization to give cross-linked filaments⁵⁵. As 1 molecule of α -actinin per 200 profilactin molecules is sufficient to cause considerable polymerization, molecules such as α -actinin may serve to mobilize actin from non-filamentous pools in cells.

Vinculin. Another protein that cross-links actin filaments is vinculin, isolated originally from smooth muscle⁴⁵. Immunofluorescence showed that it is concentrated at the termini of stress fibres near the leading edges of cells and in the region of the adhesion plaques of cultured fibroblasts, but not in surface ruffles^{45,56}. Because detergents are not required to solubilize vinculin, it does not seem to be an integral membrane protein. However, as it is located where microfilaments terminate close to membranes⁵⁷, it may participate in anchoring microfilaments together at specific membrane sites. In contrast to α -actinin, vinculin reduces the low-shear viscosity of F-actin solutions^{48,50}, apparently by forming bundles of tightly packed actin filaments in a paracrystalline manner^{48,58}. This 'bundling' role of vinculin is not universally accepted⁵⁰. Furthermore, vinculin decreases the rate of polymerization of G-actin, which again contrasts with the behaviour of α -actinins.

One feature which vinculin shares with α -actinin is the Ca^{2+} -sensitivity of its interaction with actin. Vinculin isolated from muscle shows no sensitivity to Ca^{2+} ion concentration, while the protein purified from HeLa cells (in two separable forms) causes a decrease in viscosity only at Ca^{2+} concentration $> 1 \mu\text{M}$ (ref. 50). More information is needed to establish the function of this protein.

Fascin and fimbrin. In 1975, Kane⁵⁹ demonstrated that isotonic extracts of sea urchin eggs formed gels at low Ca^{2+} concentrations. One protein responsible is fascin (M_r , 58,000), which binds to actin to form needle-like structures composed of parallel arrays of filaments with cross-bridges at 11 nm intervals⁶⁰. This interaction is not sensitive to Ca^{2+} ions. Addition of another component of M_r 220,000 from the original gel causes aggregation of the needles into a three-dimensional network. Immunofluorescence studies show that fascin is localized in the microvillar cores which form after fertilization of the sea urchin eggs⁶¹. Thus it may be involved directly in microvillus organization and morphogenesis.

Intestinal brush border microvilli also contain proteins which stabilize the core microfilament bundles. Antibodies to one of these proteins, fimbrin, also stain membrane ruffles, microspikes and microvilli in several types of cultured cells⁶². Fimbrin binds F-actin at a saturation level of 1 fimbrin per 2–3 actin monomers, to form tightly packed bundles⁶³. In native microvilli there is a second actin cross-linking protein, villin, which will be considered in detail later. These two proteins act in a concerted manner to form filament bundles when mixed together with actin at molar ratios approximating those in isolated core bundles, that is, villin/fimbrin/actin monomers = 1 : 1 : 10.

Actin severing, capping and depolymerizing proteins

β -Actinin and gelsolins. The earliest evidence for a protein that reduced the viscosity of actin filament solutions came from the work of Maruyama. He subsequently purified this protein,

which he called β -actinin, from a muscle extract. Gel filtration suggested a native M_r of $\sim 70,000$ and two bands of M_r 34,000 and 37,000 were obtained on SDS gels. β -actinin was shown to inhibit actin filament growth and re-annealing but to facilitate polymerization of G-actin⁶⁴. These two apparently contradictory properties will become a recurring theme (see below). Subsequently, several proteins that disrupt actin gels have been found in both mammalian and amoeboid systems.

Yin and Stossel⁶⁵ identified gelsolin, a calcium-dependent regulatory protein that solated actin-containing gel extracts from macrophage cytoplasm. Gelsolin is a globular monomeric protein of M_r 91,000, with two calcium-specific binding sites ($K_d = 9.2 \times 10^{-7}$ M)⁶⁶. Evidence obtained from viscometry, flow birefringence and electron microscopy indicates that at low molar ratios to actin, gelsolin rapidly cleaves the actin filaments into smaller pieces⁶⁷. It was proposed on the basis of classical polymer theory⁶⁸ that this fragmentation would profoundly affect the network properties of actin. Wang and Bryan⁶⁹ described a platelet protein of M_r 90,000 which increased the rate of nuclei formation from G-actin and blocked monomer addition to the barbed ends of subfragment-1 (S-1)-decorated filaments.

Taken together, these observations suggest the existence of a severing protein which caps filament ends to prevent re-annealing, but is also capable of complexing two or more actin monomers in polymerizing conditions to nucleate assembly. Recent studies with macrophage gelsolin have demonstrated very clearly this dual function⁷⁰. Elongation of pre-formed actin nuclei in the presence of G-actin was slowed down, whereas polymerization was facilitated in the absence of such nuclei. Electron microscopy showed preferred growth from the pointed ends when gelsolin was added to S-1-decorated actin filament fragments. Thus gelsolin binds to both G- and F-actin and in both cases an increased number of shorter filaments is produced.

Several workers have reported the presence of an actin-depolymerizing factor in blood plasma⁷¹⁻⁷³. The purified protein has a molecular weight of 90,000 (refs 74, 75). Recent experiments suggest that this protein is very similar if not identical to gelsolin, in respect of its immunological cross-reactivity⁷⁶ and its fragmentation patterns on gel electrophoresis (B. Pope, unpublished results). Furthermore, the protein behaves like gelsolin in accelerating actin polymerization in the presence of Ca^{2+} ions (ref. 75 and H. E. Harris and A.W., unpublished results). Increasing the concentration of plasma gelsolin (within a molar ratio to actin monomers of 1:200 to 1:50) produced filaments of decreasing length and gel electrophoresis showed most of the gelsolin sedimenting with the F-actin⁷⁵. We have shown that this protein releases depolymerized actin from filaments very rapidly and in a calcium-independent manner⁷⁷. Furthermore, a complex of the plasma protein with one or two actin monomers has been identified (H. E. Harris, unpublished results). These results cannot be reconciled within a model in which the sole action of gelsolin on actin filaments is severing, followed by capping of the barbed ends. The paradox may be reconciled partly by differences in the experimental conditions and by the fact that gelsolin binds to both G- and F-actin, but it is premature to speculate a detailed model for the function of this protein.

Fragmin and related proteins. Proteins that sever actin filaments have been isolated from several amoeboid cells. Fragmin is a protein which severs actin filaments and nucleates polymerization of monomers in a calcium-sensitive manner^{78,79}. It also forms a stable dimer with G-actin, but this has no severing activity. Tropomyosin or heavy meromyosin binding protect against fragmentation of filaments⁸⁰. A protein having similar properties has been purified from *Dictyostelium*⁸¹ and *Acanthamoeba* contains a capping protein that prevents gel formation by binding to the barbed ends of filaments, although this does not seem to have any severing activity⁸². A much smaller depolymerizing protein has been purified from brain, but its kinetics of actin disassembly are rather different from that of gelsolin⁸³.

Villin. Villin deserves special attention because it has both actin-severing and bundling activities, combining the properties of gelsolin and α -actinin. Villin is the major actin-associated protein in the core bundles of microfilaments from microvilli of intestinal epithelial cells. These microvilli are $\sim 1-2$ μm long and each contains 20-30 actin filaments, held together as core bundles by villin and fimbrin. The bundles are attached to the inner membrane surface by cross-connections with a periodicity of 33 nm (see ref. 84 for review). Another protein (M_r 110,000)⁸⁵ forms the cross-connections to the membrane surface at 33-nm spacings, while the remaining protein component in demembrated cores is calmodulin^{86,87}.

Villin has two distinct properties in its interaction with F-actin. In the absence of Ca^{2+} ions ($<10^{-7}$ M), it cross-links actin filaments into bundles, but when the Ca^{2+} concentration is raised above 10^{-6} M, not only are the bundles dispersed, but the filaments are also fragmented. These dual properties of villin have been intensively studied⁸⁷⁻⁹³. In addition, villin binds to actin-DNase I complexes in the presence of Ca^{2+} ions but is eluted from these complexes when the calcium is removed. This provides rapid purification of the protein⁸⁹. The paradox that villin does not bind to actin-DNase in the absence of calcium but bundles F-actin is not understood, but may reflect differences in its binding properties with monomeric and polymeric actin. Complexes of villin/actin/DNase (1:2:2) have been isolated by sucrose density gradient centrifugation⁸⁷. The presence of two actin monomers bound to villin suggests that it may nucleate actin filament assembly of G-actin: both villin and villin-actin complexes accelerate the rate of filament formation^{90,91}, the villin remaining attached to the filaments at their barbed ends⁹⁰. These reactions occur only in the presence of Ca^{2+} ions (villin has one calcium-binding site, $K_d = 2.5 \times 10^{-6}$ M; ref 87). Addition of calcium chelators to villin-severed filaments causes re-annealing⁹², suggesting that villin capping activity is lost when calcium is removed.

Selective proteolysis by V-8 protease cleaves villin into a core fragment of M_r 90,000 and a headpiece of M_r 8,000⁹⁴. The core fragment behaves rather like gelsolin—it promotes assembly of G-actin and severs filaments, but no longer bundles F-actin^{90,94}. By contrast, the headpiece binds F-actin independently of Ca^{2+} concentration and competitively inhibits F-actin bundling by villin⁹⁵. It is not clear whether bundling requires a dimer of villin, but this seems unlikely because villin does not self-associate in solution⁸⁹. In addition, it is unknown how calcium ions disrupt bundling and instead cause severing and capping. If, as expected, both headpiece and core domains are involved in actin bundling, the interaction of headpiece must be modulated by Ca^{2+} ions. The severing mechanism is totally obscure; it could occur by direct intercalation into filaments or by weakened actin-actin interactions resulting from lateral attachment of villin to F-actin. The physiological function of the calcium-sensitive, actin-severing activity of villin is obscure, but may be related to the vesiculation of microvilli that occurs in nutritional stress and in other pathological conditions⁹³, and a role for the large amount of calmodulin found in microvilli may be to prevent this vesiculation from occurring during normal uptake of Ca^{2+} ions.

The actin bundles in the microvilli extend into the terminal web, where they seem to be interconnected by fibrillar structures. In addition there are interconnections to a band of actin filaments encircling the cells close to the *zonula adherens*. Myosin and tropomyosin are concentrated near the rootlets of the microvillar bundles^{96,97}, with tropomyosin and α -actinin being concentrated in the region of the *zonula adherens*⁹⁶. The terminal web also contains a complex meshwork of 10-nm filaments which may serve to maintain rigidity of the microvilli and the adherence to adjacent epithelial cells.

Regulatory mechanisms

Table 1 summarizes information about the different classes of actin-binding proteins discussed here. It is clear that Ca^{2+} ions regulate the formation of actin bundles or gel networks in two

distinct ways. Cross-links formed by α -actinins are broken when the Ca^{2+} concentration rises above micromolar levels, but rapidly reform when calcium is removed. Actin gels are also disrupted in the presence of calcium by gelsolins, but in this case, removal of the calcium does not restore the gel network immediately as re-annealing of filaments occurs at a slower rate⁶⁷. In addition, Ca^{2+} ions independently regulate myosin filament formation and acto-myosin interaction via a calmodulin-dependent myosin light-chain kinase⁹⁸.

Calcium transients can therefore cause both contraction and solation. In some situations these phenomena may occur together, with localized disruption of gels and microfilaments sliding past one another as in muscle⁵². However, much greater contractile force may be produced if the gel network contracts without solation⁹⁹. Information is now needed to establish the precise Ca^{2+} concentrations required for these different processes and the extent to which they are independent or synchronized.

Calcium ions also regulate the interaction of gelsolins and other capping proteins with both monomeric actin and with the ends of pre-existing filaments. The function of these capping activities in cells is unclear. Based on the behaviour of these proteins *in vitro*, I suggest two different properties. In conditions where steady-state treadmilling of actin may occur, capping the barbed ends of filaments would be expected to facilitate disassembly of monomers from their pointed ends¹⁸. This could lead to complete depolymerization if the monomers were complexed with profilin or taken up by other actin filaments. Alternatively, capping proteins may function, together with free monomers, to promote actin filament formation with preferred polarity and thereby control the localization of microfilaments in cells. But how could this be achieved? One possibility is that capping factors may themselves bind to membrane proteins directly.

Actin filament growth *in vivo* may occur predominantly from the barbed ends¹⁰⁰. Assembly from stabilized monomer pools can be extremely rapid: the formation of the acrosomal process in *Thyone* sperm is the best example of this¹⁰¹. How is this triggered? Current evidence suggests that a transient proton flux may initiate acrosome formation¹⁰². The state of the non-polymerized actin in *Thyone* sperm is as yet unclear, but it is likely that a profilin-like protein stabilizes the monomer pool as has been demonstrated elsewhere¹⁹⁻²¹. In platelets it has been shown that activation by thrombin is correlated with actin filament formation¹⁰³ and this process can be inhibited by cytochalasin drugs, which also prevent the

morphological changes that normally occur after platelet activation¹⁰⁴.

Our understanding of the organization and interactions of various actin-binding proteins with actin filaments in striated muscle, both during development and in active contraction, owes much to the highly organized arrangements of these proteins and their assemblies in the sarcomeres. Similar ultrastructural analysis is being carried out on the intestinal microvillus, which also is a stable and well-organized structure having relatively few components. In motile systems, however, structures like contractile rings or pseudopodia are probably too labile for X-ray analysis and may be too poorly organized for three-dimensional image reconstruction from electron micrographs. The mechanisms by which contractile elements are moved and relocated in cells, for example in cytokinesis or cytoplasmic streaming, remain a mystery: it seems probable that macromolecular structures are not moved intact, but require disassembly and reassembly. Specific organizing centres may exist, such as are known for microtubules.

Research over the past five years has identified various actin-binding proteins in a number of cell models. Here I have attempted to group these into a simple classification. Biochemical analysis is progressing rapidly, but it is important now to investigate the roles of these components in their cellular locations. One obvious approach involves microinjection¹⁰⁵, for example, using suitably labelled actin-binding proteins or antibodies to these components where they already exist in cells. In this way we may be able to characterize particular motile events in relation to specific proteins. If these motile events are triggered by transient fluxes of Ca^{2+} or other ions, then fluorescent or chromophoric reagents can be used to explore the localization and kinetics of these ion movements. In addition, the target proteins (for example, calmodulin in the case of calcium) must be identified and their subsequent activities established. One obvious regulatory mechanism involves protein phosphorylation, but direct binding by the target protein may also modulate the properties of actin-binding proteins. Knowledge of the molecular events and regulatory mechanisms involved in cell motility is a major scientific challenge and may prove of considerable medical importance.

I thank Professor J. A. Spudich and his colleagues at the Department of Structural Biology, Stanford University for much stimulating discussion during a sabbatical year. I also thank Professors K. Burridge, J. Condeelis, B. Jockusch, E. Korn, M. Mooseker, T. Pollard, T. Stossel, L. Tilney and K. Weber for providing preprints.

- Buckley, I. K. & Porter, K. R. *Protoplasma* **64**, 345-349 (1967).
- Lazarides, E. & Weber, K. *Proc. natn. Acad. Sci. U.S.A.* **71**, 2268-2272 (1974).
- Webster, R. E., Osborn, M. & Weber, K. *Exptl Cell Res.* **117**, 47-61 (1974).
- Lazarides, E. *J. supramolec. Struct.* **5**, 531-563 (1976).
- Lazarides, E. & Burridge, K. *Cell* **6**, 289-298 (1975).
- Woloszewicz, J. & Porter, K. R. *J. Cell Biol.* **82**, 114-139 (1979).
- Heuser, J. & Kirschner, M. W. *J. Cell Biol.* **86**, 212-234 (1979).
- Hitchcock, S. E. *J. Cell Biol.* **74**, 1-15 (1977).
- Korn, E. D. *Proc. natn. Acad. Sci. U.S.A.* **75**, 588-599 (1978).
- Taylor, D. L. & Condeelis, J. S. *Int. Rev. Cytol.* **56**, 57-144 (1979).
- Uyemura, D. G. & Spudich, J. A. in *Biological Regulation and Development* Vol. 2 (ed. Goldberger, R. F.) 317-338 (Plenum, New York, 1980).
- Vandekerckhove, J. & Weber, K. *J. molec. Biol.* (1978) **126**, 783-802 (1978).
- Oosawa, F. & Asakura, S. *Thermodynamics of the Polymerization of Proteins* (Academic, London, 1975).
- Wegner, A. *J. molec. Biol.* **108**, 139-150 (1976).
- Pollard, T. D. & Mooseker, M. S. *J. Cell Biol.* **88**, 654-659 (1981).
- Margolis, R. L. & Wilson, L. *Cell* **13**, 1-8 (1978).
- Pardee, J. D., Simpson, P. A., Stryer, L. & Spudich, J. A. *J. Cell Biol.* (in the press).
- Kirschner, M. W. *J. Cell Biol.* **86**, 330-334 (1980).
- Carlsson, L., Nystrom, L.-E., Sundkvist, I., Markey, F. & Lindberg, U. *J. molec. Biol.* **115**, 465-483 (1977).
- Harris, H. E. & Weeds, A. G. *FEBS Lett.* **90**, 84-88 (1978).
- Reichstein, E. & Korn, E. D. *J. biol. Chem.* **254**, 6174-6179 (1979).
- Lazarides, E. & Linberg, U. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4742-4746 (1974).
- Brenner, S. L. & Korn, E. D. *J. biol. Chem.* **254**, 9982-9985 (1979).
- Flanagan, M. D. & Lin, S. *J. biol. Chem.* **255**, 835-838 (1980).
- McLean-Fletcher, S. & Pollard, T. D. *Cell* **20**, 329-341 (1980).
- Brown, S. S. & Spudich, J. A. *J. Cell Biol.* **88**, 487-491 (1981).
- Wang, K., Ash, J. F. & Singer, S. J. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4483-4486 (1975).
- Hartwig, J. H. & Stossel, T. D. *J. biol. Chem.* **250**, 5699-5705 (1975).
- Tyler, J. M., Anderson, J. M. & Branton, D. *J. Cell Biol.* **85**, 489-495 (1980).
- Brotschi, E. A., Hartwig, J. H. & Stossel, T. P. *J. biol. Chem.* **253**, 8988-8993 (1978).
- Hartwig, J. H. & Stossel, T. P. *J. molec. Biol.* **145**, 563-581 (1981).
- Stendahl, O. I. & Stossel, T. P. *Biochem. biophys. Res. Commun.* **92**, 675-681 (1980).
- Hartwig, J. H., Tyler, J. & Stossel, T. P. *J. Cell Biol.* **87**, 841-848 (1981).
- Valerius, N. H., Stendahl, O. I., Hartwig, J. H. & Stossel, T. P. *Cell* **24**, 195-202 (1981).
- Nunnally, M. H., D'Angelo, J. M. & Craig, S. W. *J. Cell Biol.* **87**, 219-226 (1980).
- Steck, T. L. *J. Cell Biol.* **62**, 1-19 (1974).
- Branton, D., Cohen, C. M. & Tyler, J. *Cell* **24**, 24-32 (1981).
- Gratzer, W. S. *Biochem. J.* **198**, 1-8 (1981).
- Condeelis, J., Salisbury, J. & Fujiwara, K. *Nature* **292**, 161-163 (1981).
- Condeelis, J. & Geosits, S. *J. Cell Biol.* (in the press).
- Ebashi, S., Ebashi, F. & Maruyama, K. *Nature* **203**, 645-6 (1964).
- Granger, B. L. & Lazarides, E. *Cell* **15**, 1253-1268 (1978).
- Suzuki, A. *et al. J. Biol. Chem.* **251**, 6860-6870 (1976).
- Zeece, M. G., Robson, R. M. & Betschelt, P. J. *Biochim. biophys. Acta* **581**, 365-370 (1979).
- Geiger, B. *Cell* **18**, 193-205 (1979).
- Geiger, B., Tokuyasu, K. T. & Singer, S. J. *Proc. natn. Acad.* (1979).
- Burridge, K. & McCullough, L. J. *J. supramolec. Struct.* **13**, 53-65 (1980).
- Jockusch, B. M. & Isenberg, G. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3005-3009 (1981).
- Burridge, K. & Feramisco, J. R. *Nature* **294**, 565-567 (1981).
- Burridge, K. & Feramisco, J. R. *Cold Spring Harb. Symp. quant. Biol.* **46** (in the press).
- Mimura, M. & Asano, A. *Nature* **282**, 44-48 (1979).
- Hellewell, S. B. & Taylor, D. L. *J. Cell Biol.* **83**, 633-648 (1979).
- Condeelis, J. *Int. Cell Biol.* (ed. Schweiger, H.) 306-320 (Springer, Berlin, 1981).
- Pollard, T. D. *J. biol. Chem.* **256**, 7666-7670 (1981).
- Blikstad, I., Eriksson, S. & Carlsson, L. *Eur. J. Biochem.* **109**, 317-323 (1980).
- Burridge, K. & Feramisco, J. R. *Cell* **19**, 587-595 (1980).
- Avnur, Z. & Geiger, B. *J. molec. Biol.* **153**, 361-379 (1981).
- Jockusch, B. M. & Isenberg, G. *Cold Spring Harb. Symp. quant. Biol.* **46** (in the press).
- Kane, R. E. *J. Cell Biol.* **66**, 305-315 (1975).
- Bryan, J. & Kane, R. E. *J. molec. Biol.* 207-224 (1978).
- Otto, J. J., Kane, R. E. & Bryan, J. *Cell Motility* **1**, 31-40 (1980).
- Bretscher, A. & Weber, K. *J. Cell Biol.* **86**, 335-340 (1980).
- Glenney, J. R., Kaulfus, P., Matsudaira, P. & Weber, K. *J. biol. Chem.* **256**, 9283-9288 (1981).
- Maruyama, K. *et al. J. Biochem.* **81**, 215-232 (1977).
- Yin, H. L. & Stossel, T. P. *Nature* **281**, 583-586 (1979).
- Yin, H. L. & Stossel, T. P. *J. biol. Chem.* **255**, 9494-9493 (1980).
- Yin, H. L., Zaner, K. S. & Stossel, T. P. *J. biol. Chem.* **255**, 9494-9500 (1980).

68. Flory, P. J. *Principles of Polymer Chemistry* (Cornell University Press, New York, 1953).
69. Wang, L.-L. & Bryan, J. *Cell* **25**, 637-649 (1981).
70. Yin, H. L., Maruyama, K. & Stossel, T. P. *J. biol. Chem.* **256**, 9693-9697 (1981).
71. Chaponnier, C., Borgia, R., Rungger-Brandle, E., Weil, R. & Gabbiani, C. *Experientia* **35**, 1039-1041 (1979).
72. Norberg, R., Thorstensson, R., Utter, G. & Fragaes, A. *Aur. J. Biochem.* **100**, 575-583 (1979).
73. Harris, H. E., Bamburg, J. R. & Weeds, A. G. *FEBS Lett.* **121**, 175-177 (1980).
74. Harris, H. E. & Gooch, J. *FEBS Lett.* **123**, 49-53 (1981).
75. Harris, D. A. & Schwartz, J. H. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6798-6802 (1981).
76. Yin, H. L., Albrecht, J. H. & Fattoum, A. *J. Cell Biol.* **91**, 901-906 (1981).
77. Harris, H. E., Bamburg, J. R., Bernstein, B. W. & Weeds, A. G. *Analyt. Biochem.* **119**, 102-114 (1982).
78. Hasegawa, T., Takahashi, S., Hayashi, H. & Hatano, S. *Biochemistry* **19**, 2677-2683 (1980).
79. Hinssen, H. *Eur. J. Cell Biol.* **23**, 225-233 (1981).
80. Hinssen, H. *Eur. J. Cell Biol.* **23**, 234-240 (1981).
81. Brown, S. S., Yamamoto, K. & Spudich, J. A. *J. Cell Biol.* (in the press).
82. Isenberg, G., Aebi, U. & Pollard, T. D. *Nature* **288**, 455-459 (1980).
83. Bamburg, J. R., Harris, H. E. & Weeds, A. G. *FEBS Lett.* **121**, 178-182 (1980).
84. Mooseker, M. S. & Tilney, L. G. *J. Cell Biol.* **67**, 725-743 (1975).
85. Matsudaira, P. T. & Burgess, D. R. *J. Cell Biol.* **83**, 667-673 (1979).
86. Howe, C. L., Mooseker, M. S. & Graves, T. A. *J. Cell Biol.* **86**, 916-923 (1980).
87. Glenney, J. R., Bretscher, A. & Weber, K. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6458-6462 (1980).
88. Bretscher, A. & Weber, K. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2321-2325 (1979).
89. Bretscher, A. & Weber, K. *Cell* **20**, 839-847 (1980).
90. Glenney, J. R., Kaulfus, P. & Weber, K. *Cell* **24**, 471-480 (1981).
91. Craig, S. W. & Powell, L. D. *Cell* **22**, 739-746 (1980).
92. Mooseker, M. S., Graves, T. A., Wharton, K. A., Falco, N. & Howe, C. L. *J. Cell Biol.* **87**, 809-822 (1980).
93. Matsudaira, P. T. & Burgess, D. R. *J. Cell Biol.* (in the press).
94. Glenney, J. R. & Weber, K. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2810-2814 (1981).
95. Glenney, J. R., Geisler, N., Kaulfus, P. & Weber, K. *J. biol. Chem.* **256**, 8156-8161 (1981).
96. Drenckhahn, D. & Groschel-Stewart, U. *J. Cell Biol.* **86**, 475-482 (1980).
97. Herman, I.-M. & Pollard, T. D. *J. Cell Biol.* **88**, 346-351 (1981).
98. Kendrick-Jones, J. & Scholey, J. *J. Muscle Res. Cell Motility* **2**, 347-372 (1981).
99. Stossel, T. P., Hartwig, J. H., Yin, H. L., Zaner, K. S. & Stendahl, O. *Cold Spring Harb. Symp. quant. Biol.* **46** (in the press).
100. Tilney, L. G., Binder, E. M. & DeRosier, D. J. *J. Cell Biol.* **90**, 485-494 (1981).
101. Tilney, L. G., Hatano, S., Ishikawa, H. & Mooseker, M. S. *J. Cell Biol.* **59**, 109-126 (1973).
102. Tilney, L. G., Kiehart, D. P., Sardet, C. & Tilney, M. J. *J. Cell Biol.* **77**, 536-550 (1978).
103. Carlsson, L., Markey, F., Blikstad, I., Persson, T. & Lindberg, U. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6376-6380 (1979).
104. Casella, J. F., Flanagan, M. D. & Lin, S. *Nature* **293**, 302-305 (1981).
105. Taylor, D. L. & Wang, Y.-L. *Nature* **284**, 405-410 (1980).

ARTICLES

UV radiation from the young Sun and oxygen and ozone levels in the prebiological palaeoatmosphere

V. M. Canuto*, J. S. Levine†, T. R. Augustsson† & C. L. Imhoff‡

* NASA, Goddard Institute for Space Studies, New York, New York 10025, USA

† NASA, Langley Research Center, Hampton, Virginia 23665, USA

‡ C.S.C., NASA, Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

UV measurements of young T-Tauri stars, resembling the Sun at an age of a few million years, have recently been made with the International Ultraviolet Explorer. They indicate that young stars emit up to 10^4 times more UV than the present Sun. The implications for the origin and evolution of O_2 and O_3 in the prebiological palaeoatmosphere are presented here. The results of photochemical calculations indicate that the O_2 surface mixing ratio was a factor 10^4 – 10^6 times greater than the standard value of 10^{-15} . This new value reconciles the simultaneous existence of oxidized iron and reduced uranium.

A RELIABLE picture of the environment of the early Earth depends critically on the correct quantification of the amount of solar radiation impinging on the Earth. Standard stellar evolutionary models predict that 4,500 Myr ago, the Sun's luminosity was lower than today by ~30% (ref. 1). This has created the 'dim Sun paradox' because such dimming translates into a decrease of the Earth's effective temperature, T_* , of ~8%, sufficient to keep T_* below the freezing point of seawater for ~2,000 Myr (ref. 2). On the other hand, there is evidence of liquid water as early as 3,500 Myr ago², so a paradox arises. Explanations proposed to solve the problem include either increasing the greenhouse efficiency²⁻⁴ or varying the Earth's early albedo⁵. In either case, one makes inferences about the

physical parameters characterizing the palaeoatmosphere using astronomical data.

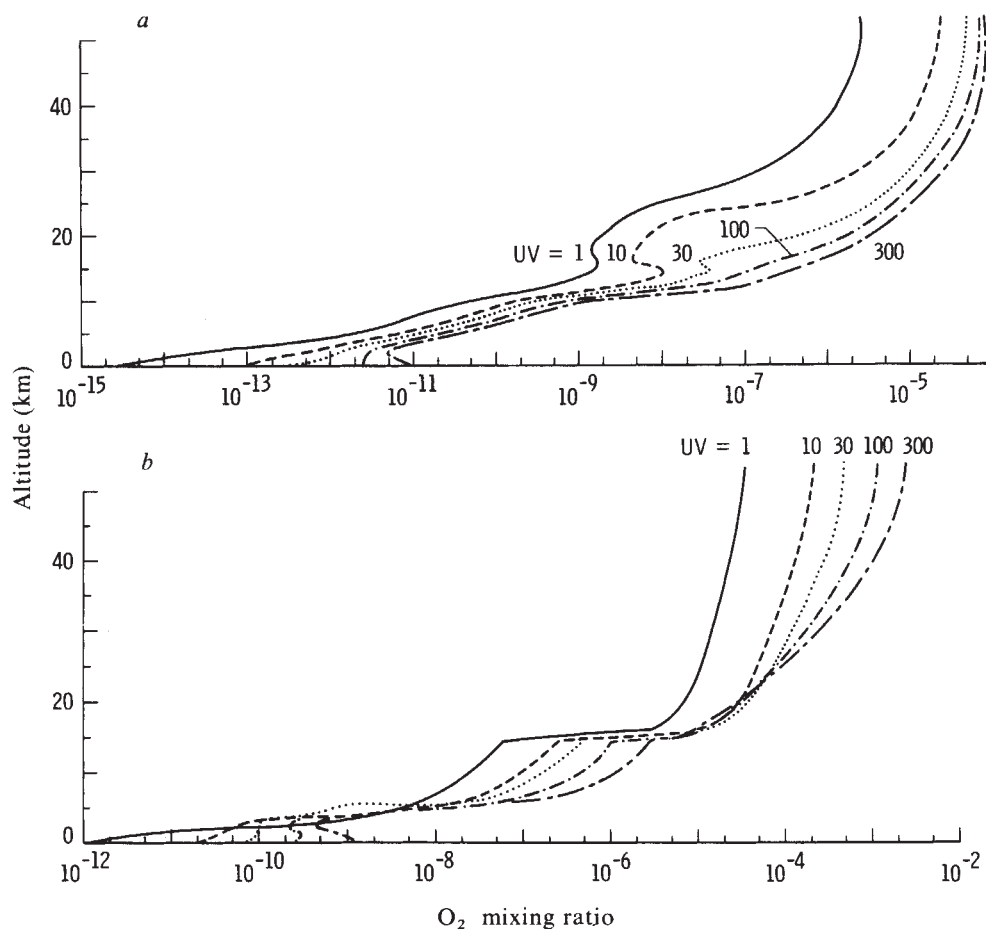
We consider here another astronomical input, the UV radiation from the Sun which had a great impact on the early atmosphere of the Earth⁶⁻¹². In fact, UV radiation initiated the photochemical processes that led to the formation of oxygen (O_2) and ozone (O_3) in the prebiological palaeoatmosphere. By virtue of its strong absorption in the UV (200–300 nm), ozone protects life at the surface of the Earth from this lethal radiation. Hence, the evolution of ozone, and its variation over geological time had very important implications for the biological evolution of primitive organisms⁶⁻¹². An accurate quantification of the levels of oxygen and ozone in the prebiological

Table 1 Ratios of stellar to solar UV line fluxes intercepted at the distance of the Earth

λ (Å)	Identity	T Tau	DR Tau	RW Aur	GW Ori	CoD-35°10525	RU Lup	S CrA
1,240	N v	2.5×10^5		$<4.0 \times 10^3$	2.3×10^5	6.1×10^4	2.4×10^3	
1,304	O I, S I	1.3×10^5		9.4×10^3	9.9×10^4	3.0×10^4	6.5×10^3	
1,335	C II	7.8×10^4		9.6×10^2	4.3×10^4	1.7×10^4	1.9×10^4	
1,400	Si IV	2.2×10^5	5.5×10^4	1.2×10^4	2.3×10^5	2.0×10^4	3.3×10^4	2.4×10^4
1,550	C IV	1.4×10^5	8.8×10^4	3.2×10^3	1.5×10^5	2.9×10^4	1.2×10^4	8.1×10^3
1,640	He II	1.4×10^5	5.9×10^4	$\leq 1.5 \times 10^3$	2.0×10^5	2.6×10^4	5.8×10^3	$\leq 7.7 \times 10^3$
1,813	Si II, S I	2.9×10^4	1.9×10^4		1.1×10^4	2.8×10^3	5.1×10^3	5.6×10^3

The radii of the stars have been taken as (in solar radii): 6.8, 4.7, 3.1, 8.4, 3.4, 2.6, and 3.2, respectively.

Fig. 1 Vertical distribution of O_2 in the prebiological palaeoatmosphere for levels of solar UV from 1 to 300 times the present value for a CO_2 level of: *a*, 280 p.p.m.v.; *b*, 28,000 p.p.m.v.



palaeoatmosphere and the level of UV radiation reaching the surface of the Earth are therefore important when considering biological evolution on Earth.

The stellar evolutionary models cannot be used to predict the amount of UV generated by the young Sun and one must, therefore, rely on direct observational data. Because until recently such data were available only for the present Sun, estimates for the past UV radiation were based either on the assumption that the lower solar luminosity also implied a lower UV component or else it was assumed that the UV radiation was the same in the past as it is today⁶⁻¹².

Recent observational data gathered through the International Ultraviolet Explorer (IUE) have, however, indicated that neither approach is correct. Young stars have been found to emit considerably more UV radiation than the present Sun.

The T-Tauri stars are very young stars with masses roughly equal to the Sun's. As such, they are considered to resemble the Sun at an age of a few million years. Several T-Tauri stars have recently been studied in the UV with IUE¹³⁻¹⁷. The primary result of these observations is that the stars emit strongly in the UV. Most noticeable are the chromospheric emission lines arising from the hot (10^4 – 10^5 K) atmospheres of the stars. The surface fluxes of the lines, (defined to be the flux emitted per unit stellar surface area), have been determined to be 100–5,000 times greater among the T-Tauri stars than in the Sun¹⁸. In addition, these protostars are still contracting and thus are larger than the Sun. Thus, the UV flux intercepted at the distance of the Earth is increased by an additional factor of $(R_*/R_\odot)^2$ due to the greater surface area of the protostar. We have computed the ratio of stellar-to-solar UV flux as would be measured at the Earth's distance for seven T-Tauri stars, given in Table 1. Radii for these stars were estimated from temperatures and luminosities of the stars largely taken from ref. 19. The values range from 3 to 8 solar radii and are given

in Table 1. We find that the UV line flux from the T-Tauri stars is 10^3 – 10^4 times the present solar value.

The emission lines dominate the spectrum between 100 and 200 nm. At longer wavelengths the UV continuum of the T-Tauri stars dominates. The continuum fluxes are also enhanced by a few hundred times relative to the Sun. As we will discuss below, the wavelength intervals discussed here cover the maximum in the UV absorption cross-section of H_2O , CO_2 , O_2 , and O_3 .

Young stars which are older than the T-Tauri stars continue to emit strongly in the UV, although not at the extreme levels of the T-Tauri stars. UV emission-line surface fluxes decrease with age approximately as $t^{-0.5}$ to $t^{-0.9}$ (ref. 20). The age dependence for the UV surface fluxes resembles that of Ca II emission, determined previously by ground-based near-UV observations²²⁻²³. We have computed the enhancement of the UV flux intercepted at the distance of the Earth as a function of age for a solar-type star, using the observational relationship reported in ref. 20 corrected for the changing radius of the contracting star. Radii at specific ages for solar mass stars were estimated from temperatures and luminosities in theoretical H-R diagrams¹⁹. The results are given in Table 2. We find that, even after the star has reached the main sequence (5×10^7 yr), begins

Table 2 Stellar UV flux as a function of age

Age (yr)	UV enhancement
10^6	10^4
10^7	500
5×10^7	100
10^8	32
5×10^8	8
10^9	4
5×10^9	1

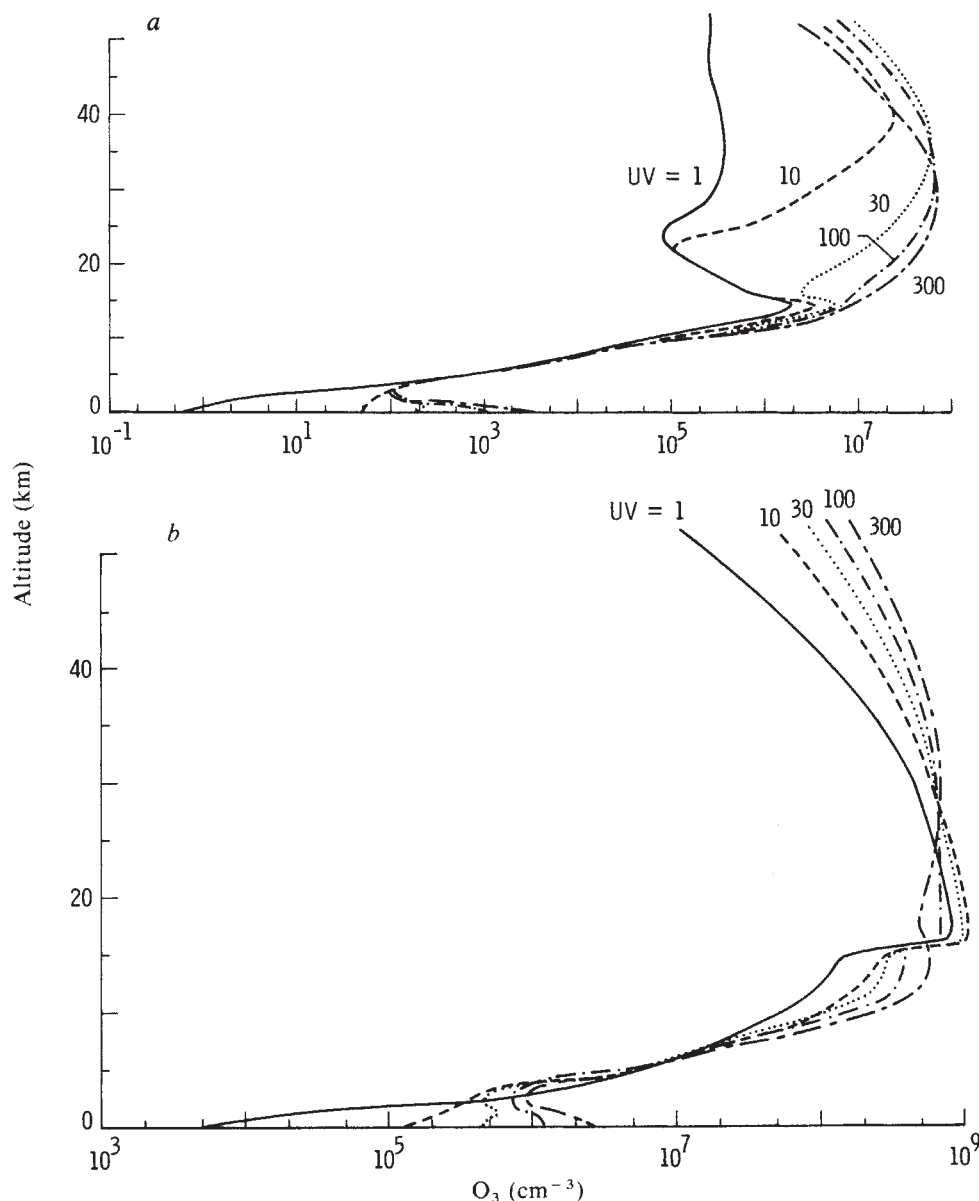


Fig. 2 Vertical distribution of O_3 in the prebiological palaeoatmosphere for levels of solar UV from 1 to 300 times the present value for a CO_2 level of: *a*, 280 p.p.m.v.; *b*, 28,000 p.p.m.v.

hydrogen burning and becomes stable, it emits about 100 times the UV radiation presently observed in the Sun.

There is no reason to believe that the youth of the Sun differed from that indicated by the T-Tauri and other young stars. We, therefore, conclude that the UV output of the young Sun was 10–1,000 times its present value. The high level of UV radiation thus indicated must have a profound effect on our understanding of the early atmosphere of the Earth.

Photochemistry of oxygen and ozone

Enhanced levels of UV radiation that may have been emitted from the Sun in its early stages may have had some very important implications for the origin and evolution of the early atmosphere, particularly for prebiological levels of oxygen (O_2) and ozone (O_3). In the prebiological palaeoatmosphere the formation of O_2 was initiated by solar UV radiation, that is, the photolysis of water vapour (H_2O), accompanied by the exospheric escape of atomic hydrogen (H), and the photolysis of carbon dioxide (CO_2). The photochemical reactions governing these processes are:



and,



Table 3 Photodissociation reactions and diurnally-averaged rates at 53.5 km for a latitude of 30° (equinoctial conditions) (for present atmosphere)

No.	Reaction	Destruction ($\text{cm}^{-3} \text{ s}^{-1}$)
J1	$O_2 + h\nu \rightarrow O + O$	1.41×10^6
J2	$O_3 + h\nu \rightarrow O + O_2$	4.40×10^6
J3	$O_3 + h\nu \rightarrow O(^1D) + O_2$	7.07×10^7
J4	$H_2O + h\nu \rightarrow OH + H$	5.32×10^2
J5	$N_2O + h\nu \rightarrow O(^1D) + N_2$	1.29×10^2
J6	$HNO_3 + h\nu \rightarrow OH + NO_2$	2.63×10^0
J7	$NO_2 + h\nu \rightarrow O + NO$	1.33×10^4
J8	$H_2O_2 + h\nu \rightarrow OH + OH$	2.52×10^2
J9	$HNO_2 + h\nu \rightarrow OH + NO$	4.31×10^1
J10	$NO_3 + h\nu \rightarrow NO + O_2$	7.85×10^{-1}
J11	$NO_3 + h\nu \rightarrow NO_2 + O$	7.85×10^0
J12	$N_2O_5 + h\nu \rightarrow NO_2 + NO_3$	5.21×10^{-4}
J13	$HCl + h\nu \rightarrow H + Cl$	1.47×10^0
J14	$ClO_2 + h\nu \rightarrow ClO + O$	6.41×10^{-3}
J15	$ClO + h\nu \rightarrow Cl + O$	1.45×10^2
J16	$Cl_2 + h\nu \rightarrow Cl + Cl$	1.43×10^{-6}
J17	$ClNO_2 + h\nu \rightarrow ClO + NO_2$	4.77×10^{-4}
J18	$CCl_4 + h\nu \rightarrow 2Cl + \text{products}$	1.27×10^{-5}
J19	$CH_3Cl + h\nu \rightarrow CH_3 + Cl$	1.09×10^{-2}
J20	$CH_3O + h\nu \rightarrow H + HCO$	4.00×10^1
J21	$CH_3O + h\nu \rightarrow H_2 + CO$	4.15×10^1
J22	$CH_3CCl_3 + h\nu \rightarrow Cl + \text{products}$	1.20×10^{-5}
J23	$CH_3OOH + h\nu \rightarrow CH_3O + OH$	1.82×10^0
J24	$NO + h\nu \rightarrow N + O$	9.12×10^1
J25	$CO_2 + h\nu \rightarrow CO + O$	4.51×10^{-1}

Table 4 Chemical reactions

No.	Reaction	Rate constant (cm ³ s ⁻¹ or cm ⁶ s ⁻¹)
1	O + O ₂ + M → O ₃ + M	1.1 × 10 ⁻³⁴ exp(510/T)
2	O + O ₂ → 2O ₂	1.5 × 10 ⁻¹¹ exp(-2,218/T)
3	O(¹ D) + O ₂ → O + O ₂	3.2 × 10 ⁻¹¹ exp(67/T)
4	O(¹ D) + N ₂ → O + N ₂	1.8 × 10 ⁻¹¹ exp(107/T)
5	N ₂ O + O(¹ D) → 2NO	6.6 × 10 ⁻¹¹
6	N ₂ O + O(¹ D) → N ₂ + O ₂	5.1 × 10 ⁻¹¹
7	NO + O + M → NO ₂ + M	1.6 × 10 ⁻³² exp(584/T)
8	NO + O ₂ → NO ₂ + O	2.3 × 10 ⁻¹² exp(-1,450/T)
9	NO ₂ + O → O ₂ + NO	9.3 × 10 ⁻¹²
10	NO ₂ + O ₃ → NO ₃ + O ₂	1.2 × 10 ⁻¹³ exp(-2,450/T)
11	NO + HO ₂ → NO ₂ + OH	3.5 × 10 ⁻¹² exp(250/T)
12	NO ₂ + OH + M → HNO ₃ + M	*
13	HNO ₃ + OH → NO ₃ + H ₂ O	1.5 × 10 ⁻¹⁴ exp(650/T)
14	H ₂ O + O(¹ D) → 2OH	2.2 × 10 ⁻¹⁰
15	H + O ₂ + M → HO ₂ + M	2.1 × 10 ⁻³² exp(290/T)
16	H + O ₃ → OH + O ₂	1.4 × 10 ⁻¹⁰ exp(-470/T)
17	OH + O → H + O ₂	2.3 × 10 ⁻¹¹ exp(110/T)
18	OH + O ₂ → HO ₂ + O	1.6 × 10 ⁻¹² exp(-940/T)
19	OH + OH → H ₂ O + O	4.5 × 10 ⁻¹² exp(-275/T)
20	HO ₂ + O → OH + O ₂	4.0 × 10 ⁻¹¹
21	HO ₂ + O ₃ → OH + 2O ₂	1.1 × 10 ⁻¹⁴ exp(-580/T)
22	HO ₂ + OH → H ₂ O + O ₂	4.0 × 10 ⁻¹¹
23	HO ₂ + HO ₂ → H ₂ O ₂ + O ₂	2.5 × 10 ⁻¹²
24	H ₂ O ₂ + OH → HO ₂ + H ₂ O	2.7 × 10 ⁻¹² exp(-145/T)
25	OH + NO + M → HNO ₂ + M	*
26	NO + NO ₃ → 2NO ₂	2.0 × 10 ⁻¹¹
27	O(¹ D) + N ₂ + M → N ₂ O + M	3.5 × 10 ⁻³⁷
28	O(¹ D) + H ₂ → OH + H	9.9 × 10 ⁻¹¹
29	O(¹ D) + CH ₄ → OH + CH ₃	1.4 × 10 ⁻¹⁰
30	NO ₂ + O + M → NO ₃ + M	1.0 × 10 ⁻³¹
31	NO ₂ + NO ₃ → N ₂ O ₅	1.5 × 10 ⁻¹³ exp(861/T)
32	N ₂ O ₅ → NO ₂ + NO ₃	1.2 × 10 ⁻¹⁴ exp(-10,319/T)
33	N ₂ O ₅ + H ₂ O → 2HNO ₃	1.0 × 10 ⁻²⁰
34	N ₂ O ₅ + O → 2NO ₂ + O ₂	3.0 × 10 ⁻¹⁶
35	N + NO ₂ → N ₂ O + O	2.1 × 10 ⁻¹¹ exp(-800/T)
36	N + O ₂ → NO + O	4.4 × 10 ⁻¹² exp(-3,220/T)
37	N + NO → N ₂ + O	3.4 × 10 ⁻¹¹
38	N + O ₃ → NO + O ₂	1.0 × 10 ⁻¹⁵
39	Cl + O ₃ → ClO + O ₂	2.8 × 10 ⁻¹¹ exp(-257/T)
40	ClO + O → Cl + O ₂	7.7 × 10 ⁻¹¹ exp(-130/T)
41	ClO + NO → Cl + NO ₂	6.5 × 10 ⁻¹² exp(280/T)
42	Cl + CH ₄ → HCl + CH ₃	9.6 × 10 ⁻¹² exp(-1,350/T)
43	Cl + H ₂ → HCl + H	3.5 × 10 ⁻¹¹ exp(-2,290/T)
44	Cl + HO ₂ → HCl + O ₂	4.8 × 10 ⁻¹¹
45	Cl + H ₂ O ₂ → HCl + HO ₂	1.1 × 10 ⁻¹² exp(-980/T)
46	Cl + HNO ₃ → HCl + NO ₃	1.0 × 10 ⁻¹¹ exp(-2,170/T)
47	Cl + CH ₃ O → HCl + HCO	9.2 × 10 ⁻¹³ exp(-68/T)
48	HCl + OH → Cl + H ₂ O	2.8 × 10 ⁻¹² exp(-425/T)
49	HCl + O → Cl + OH	1.1 × 10 ⁻¹¹ exp(-3,370/T)
50	Cl + O ₂ + M → ClO ₂ + M	*
51	ClO ₂ + M → Cl + O ₂ + M	2.7 × 10 ⁻⁹ exp(-2,650/T)
52	Cl + ClO ₂ → 2ClO	8.0 × 10 ⁻¹²
53	Cl + ClO ₂ → Cl ₂ + O ₂	1.4 × 10 ⁻¹⁰
54	OH + OH + M → H ₂ O ₂ + M	*
55	H ₂ O ₂ + O → OH + HO ₂	2.8 × 10 ⁻¹² exp(-2,125/T)
56	OH + CH ₄ → CH ₃ + H ₂ O	2.4 × 10 ⁻¹² exp(-1,710/T)
57	ClO + NO ₂ + M → ClONO ₂ + M	*
58	O + ClONO ₂ → ClO + NO ₂	3.0 × 10 ⁻¹² exp(808/T)
59	O(¹ D) + HCl → OH + Cl	1.4 × 10 ⁻¹⁰
60	H ₂ + OH → H ₂ O + H	1.2 × 10 ⁻¹¹ exp(-2,200/T)
61	CH ₃ + O ₂ + M → CH ₃ O ₂ + M	*
62	CH ₃ O ₂ + HO ₂ → CH ₃ OOH + O ₂	7.7 × 10 ⁻¹⁴ exp(1,300/T)
63	CH ₃ O ₂ + NO → CH ₃ O + NO ₂	7.4 × 10 ⁻¹²
64	CH ₃ O + O ₂ → CH ₃ O ₂ + O	9.2 × 10 ⁻¹³ exp(-2,200/T)
65	CH ₃ O + OH → HCO + H ₂ O	1.0 × 10 ⁻¹¹
66	CH ₃ O → OH + HCO	3.0 × 10 ⁻¹¹ exp(-1,550/T)
67	HCO + O ₂ → CO + HO ₂	5.0 × 10 ⁻¹²
68	CO + OH → H + CO ₂	*
69	CH ₃ Cl + OH → Cl + products	1.8 × 10 ⁻¹² exp(-1,112/T)
70	CH ₃ OOH + OH → CH ₃ O ₂ + H ₂ O	2.1 × 10 ⁻¹² exp(-145/T)
71	OH + CH ₃ CCl ₃ → Cl + products	5.4 × 10 ⁻¹² exp(-1,820/T)
72	O + O + M → O ₂ + M	2.8 × 10 ⁻³⁴ exp(710/T)
73	H + H + M → H ₂ + M	8.3 × 10 ⁻³³
74	H ₂ + O → OH + H	3.0 × 10 ⁻¹⁴ exp(-4,480/T)
75	H + HO ₂ → O ₂ + H ₂	4.7 × 10 ⁻¹¹ (×0.29)
76	H + HO ₂ → H ₂ O + O	4.7 × 10 ⁻¹¹ (×0.02)
77	H + HO ₂ → OH + OH	4.7 × 10 ⁻¹¹ (×0.69)

* See ref. 34.

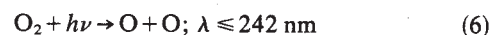
Reaction (1) leads to the production of oxygen atoms (O) through:



The oxygen atoms formed in reactions (2) and (3) form molecular oxygen (O₂) through the following reactions:



The destruction of O₂ is controlled by its photolysis and reaction with molecular hydrogen (H₂) resulting from volcanic emissions:



and



(See reactions (19) and (74) in Table 4.)

Closely coupled to the origin and evolution of O₂ is the origin and evolution of O₃. O₃ is photochemically produced through the reaction:

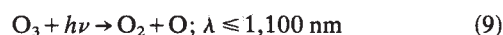
Table 5 Column density of O₃ in the prebiological palaeoatmosphere for various combinations of atmospheric CO₂ and solar UV flux

Solar UV flux	CO ₂	O ₃ column (cm ⁻²)
1*	1†	1.72 × 10 ¹²
	10	2.49 × 10 ¹⁴
	100	1.38 × 10 ¹⁵
10	1	3.94 × 10 ¹³
	10	5.36 × 10 ¹⁴
	100	1.95 × 10 ¹⁵
30	1	1.50 × 10 ¹⁴
	10	7.81 × 10 ¹⁴
	100	2.06 × 10 ¹⁵
100	1	1.76 × 10 ¹⁴
	10	8.15 × 10 ¹⁴
	100	2.04 × 10 ¹⁵
300	1	1.52 × 10 ¹²
	10	9.41 × 10 ¹⁴
	100	2.19 × 10 ¹⁵

* Present value of solar UV flux²⁷. Other values represent multiples of present value.

† Pre-industrial level of CO₂ (p.p.m.v.). Other values represent multiples of this value.

O₃ is photochemically destroyed through photolysis and reaction with O:



and



In addition, O₃ is photochemically destroyed by a series of catalytic cycles involving the oxides of nitrogen (NO + NO₂), hydrogen (OH + HO₂), and chlorine (Cl + ClO).

To study the effects of enhanced solar UV radiation on the levels of O₂ and O₃ in the prebiological palaeoatmosphere, the photochemical model of the palaeoatmosphere of Levine *et al.*²⁴ was modified. This model which was originally developed to study the effect of anthropogenic perturbations on future

levels of O_3 , has been used to study the origin and evolution of O_3 for specified levels of palaeoatmospheric O_2 ranging from 10^{-4} of the present atmospheric level (PAL) to the present atmospheric level (1 PAL)²⁴. The model was modified such that O_2 , CO_2 , and H_2 which were previously specified as input parameters are now considered as chemically-active species, whose vertical profiles are calculated using coupled species continuity-flux equations, which contain both chemical production and loss terms and vertical eddy transport²⁵. The inclusion of O_2 , CO_2 , and H_2 as chemically-active species makes it a total of 31 species whose vertical profiles are solved simultaneously by coupled species continuity-flux equations. The other calculated species are: O , N_2O , N , NO , NO_2 , NO_3 , N_2O_5 , HNO_2 , HNO_3 , H_2O , H , OH , HO_2 , H_2O_2 , CH_4 , CO , CH_3OOH , CH_2O , CH_3 , HCO , CH_3O_2 , CH_3O , Cl , ClO , HCl , CH_3Cl , and $ClNO_3$. Very short-lived species assumed to be in instantaneous photochemical equilibrium include: $O(^1D)$, ClO_2 , and Cl_2 . The model now includes 25 photochemical processes listed in Table 3, and 77 chemical reactions listed in Table 4. The US Standard Atmospheric Mid-Latitude Spring/Autumn temperature and H_2O profiles are specified in the troposphere. A primordial temperature profile that decreases linearly from the tropopause to the mesosphere is used. This profile is based on coupled photochemical-radiative equilibrium temperature calculations in the O_3 -deficient palaeoatmosphere²⁶. While H_2O is specified in the troposphere, it is calculated through the continuity-flux equation for H_2O in the stratosphere. Water soluble species, for example, HNO_3 , H_2O_2 , CH_3OOH , HCl , ClO , and $ClNO_3$ are lost throughout rainout, with a rainout coefficient of $1.0 \times 10^{-6} s^{-1}$. The model extends from the surface up to 53.5 km with the tropopause at a height of 14.5 km. Photodissociation rates are diurnally-averaged for a latitude of 30° and solar declination of 0° . The model includes the solar spectrum from 110 to 735 nm (ref. 27). Further details about the model are given in refs 24, 25.

For the present calculations, we have incorporated the boundary conditions for O_2 and H_2 given in ref. 28. O_2 is considered as a chemically-active species with a zero flux lower boundary condition and H_2 has a prescribed surface mixing ratio of 17 p.p.m.v. (ref. 28). We have studied the sensitivity of prebiological O_2 and O_3 to the level of CO_2 by performing calculations for the pre-industrial CO_2 value (280 p.p.m.v.), and for 10–100 times this value. Volatile outgassing scenarios suggest that the early atmosphere may have contained as much as three orders of magnitude more CO_2 than found in the present atmosphere^{29,30}. According to these scenarios, the bulk of the early CO_2 went into the ocean and eventually formed sedimentary carbonates. The outgassed CO_2 on both Mars and Venus had a very different fate—remaining in the atmosphere—presumably due to the lack of liquid water and life on these planets. For our calculations we have used five values for the flux of solar UV: the present value²⁷ and multiples of 10, 30, 100, and 300 times the present value.

Oxygen and ozone in the prebiological palaeoatmosphere

O_2 profiles for the pre-industrial level of CO_2 (280 p.p.m.v.) and 100 times this value are given in Fig. 1. For each CO_2 level calculations have been performed for five values of solar UV radiation ranging from 1 to 300 times the present value. The calculated O_2 profile for contemporary values of CO_2 and UV exhibits a very strong altitude dependence, with a minimum O_2 mixing ratio of $<10^{-14}$ at the surface and a maximum O_2 mixing ratio of about 2×10^{-6} above 40 km, in excellent agreement with earlier calculations²⁸. For a CO_2 value of 280 p.p.m.v. (Fig. 1), we found that increasing the UV flux from the present value to 300 times the present value resulted in increasing the surface mixing ratio of O_2 from $<10^{-14}$ to $\sim 10^{-11}$. The O_2 maximum between 40 and 50 km increased from a mixing ratio of $\sim 2 \times 10^{-6}$ to 8×10^{-5} as the solar UV flux was increased by a factor of 300. For a CO_2 value of 100 times the pre-industrial value (Fig. 1b), increasing the UV flux by a factor of 300 increased the surface mixing ratio of O_2 from 10^{-12} to 10^{-9} , and increased the O_2 maximum at ~ 50 km from $\sim 3 \times 10^{-5}$ to $\sim 2 \times 10^{-3}$. The corresponding O_3 calculations are shown in Fig. 2a (for $CO_2 = 280$ p.p.m.) and Fig. 2b (for $CO_2 = 100$ times greater). For a CO_2 value of 280 p.p.m.v. (Fig. 2a), increasing the UV flux by a factor of 300 increased the surface concentration of O_3 from 1 to $\sim 3 \times 10^3 cm^{-3}$ and increased both the height of the O_3 maximum from 14.5 to ~ 30 km and the peak concentration from 1×10^6 to $\sim 5 \times 10^7 cm^{-3}$. For the enhanced value of CO_2 (Fig. 2b), increasing the UV flux by a factor of 300 increased the surface concentration of O_3 from 5×10^3 to $2 \times 10^6 cm^{-3}$. For this level of CO_2 and for all values of enhanced UV radiation the height of O_3 maximum was found to be between 15 and 20 km with a concentration that varied between 5 and $10 \times 10^8 cm^{-3}$. The total O_3 column (molecules cm^{-2}) above the surface of the Earth for the O_3 profiles shown in Fig. 2 and for a CO_2 value of 10 times the pre-industrial value is summarized in Table 5.

Conclusion

Our calculation indicates that the surface mixing ratio of O_2 increased by a factor of 10^4 – 10^6 times the standard value of 10^{-15} (depending on the assumed value of CO_2) for enhanced level of UV radiation associated with the early Sun. Corresponding enhancement of O_3 level was also found.

These increased levels of O_2 and O_3 may have had important implications for early geology and on the origin and evolution of life. While the details of such implications are beyond the scope of this article, note that the higher value of O_2 found here falls within the lower (10^{-13}) and upper (10^{-3}) O_2 values arrived at to explain the simultaneous existence of oxidized iron and reduced uranium on the early Earth^{31–33}.

V.M.C. and J.S.L. thank Dr K. M. Towe for his helpful suggestions. C.L.I. thanks Dr M. S. Giampapa for assistance.

Received 28 December 1981; accepted 23 February 1982.

- Newman, M. J. & Rood, R. T. *Science* **198**, 1035 (1977).
- Sagan, C. & Mullen, G. *Science* **177**, 52 (1972).
- Sagan, C. *Science* **209**, 224 (1977).
- Owen, T., Cess, R. D. & Ramanathan, V. *Nature* **277**, 640 (1979).
- Henderson-Sellers, A. & Meadows, A. J. *Nature* **270**, 589 (1977).
- Berkner, L. V. & Marshall, L. C. *J. Atmos. Sci.* **22**, 225 (1965).
- Ratner, M. I. & Walker, J. C. G. *J. Atmos. Sci.* **29**, 803 (1972).
- Blake, A. J. & Carver, H. J. *J. Atmos. Sci.* **34**, 720 (1977).
- Levine, J. S., Hays, P. B. & Walker, J. C. G. *Icarus* **39**, 295 (1979).
- Kasting, J. F. & Donahue, T. M. *J. geophys. Res.* **85**, 3255 (1980).
- Levine, J. S., Boughner, R. E. & Smith, K. A. *Origins of Life* **10**, 199 (1980).
- Levine, J. S. *Origins of Life* **10**, 313 (1980).
- Appenzeller, I., Chavarría, C., Krautter, J., Mundt, R. & Wolf, B. *Astr. Astrophys.* **90**, 184 (1980).
- Appenzeller, I. & Wolf, B. *Astr. Astrophys.* **75**, 164 (1979).
- Brown, A., Jordan, C., Millar, T. J., Gondhaekar, P. & Wilson, R. *Nature* **290**, 34 (1981).
- Gahn, G. F., Kredga, K., Liseau, R. & Dravins, D. *Astr. Astrophys.* **73**, L4 (1979).
- Gondhaekar, P., Penston, M. V. & Wilson, R. *The First Year of IUE* (ed. Willis, A. J.) 109 (University College, London, 1979).
- Imhoff, C. L. & Giampapa, M. S. *Astrophys. J. Lett.* **239**, L115 (1980); in *The Universe at Ultraviolet Wavelengths: The First Two Years of IUE* (ed. Chapman, R. D.) 185 (NASA Conf. Publ. 2171, 1981); in *Proc. 2nd Cambridge Workshop on Cool Stars, Stellar Systems and the Sun* (ed. Giampapa, M. S.) (SAO Spec. Rep., in the press).
- Cohen, M. & Kuhl, L. V. *Astrophys. Suppl.* **41**, 743 (1979).
- Boesgaard, A. M. & Simon, T. in *Proc. 2nd Cambridge Workshop on Cool Stars, Stellar Systems and the Sun* (ed. Giampapa, M. S.) (SAO Spec. Rep., in the press).
- Wilson, O. C. *Astrophys. J.* **138**, 832 (1963).
- Wilson, O. C. & Skumanich, A. *Astrophys. J.* **140**, 1401 (1964).
- Skumanich, A. *Astrophys. J.* **171**, 565 (1972).
- Levine, J. S., Augustsson, T. R., Boughner, R. E., Natarajan, M. & Sacks, L. J. in *Comets and the Origin of Life* (ed. Ponnampuram, C.) 161 (Reidel, Dordrecht, 1981).
- Levine, J. S., Augustsson, T. R. & Natarajan, M. *Origins of Life* **12** (in the press).
- Levine, J. S. & Boughner, R. E. *Icarus* **39**, 310 (1979).
- Ackermann, M. in *Mesospheric Models and Related Experiments* (ed. Fiocco, G.) 149 (Reidel, Dordrecht, 1971).
- Kasting, J. F. & Walker, J. C. G. *J. geophys. Res.* **86**, 1147 (1981).
- Hart, M. H. *Icarus* **38**, 23 (1978).
- Levine, J. S. in *Comparative Planetology* (ed. Ponnampuram, C.) 165 (Academic, New York, 1978).
- Kline, C. & Bricker, O. P. *Econ. Geol.* **72**, 1457 (1977).
- Grandstaff, D. E. *Precamb. Res.* **13**, 1 (1980).
- Towe, K. M. *Nature* **274**, 657 (1978).
- Chemical Kinetic and Photochemical Data for Use in Stratospheric Modelling* JPL Publ. 81-3 (Jet Propulsion Laboratory, Pasadena, 1981).

Sr and Nd isotope geochemistry of oceanic basalts and mantle evolution

W. M. White^{*†‡} & A. W. Hofmann^{*‡§}

^{*} Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Rd NW, Washington DC 20015, USA

[†] Department of Chemistry and Geochemistry, Colorado School of Mines, Golden, Colorado 80401, USA

[‡] Max-Planck-Institut für Chemie, Postfach 3060, 6500 Mainz, FRG

Sr and Nd isotope ratios are reported for 17 mid-ocean ridge basalts and for 11 oceanic islands and island groups. Data from the Azores, Samoa and the Society Islands diverge significantly from the mantle array. These results are not explained by binary mixing of depleted and undepleted mantle reservoirs or by variable magmatic depletion of a planetary reservoir, but support mantle evolution models involving re-injection of crust material into the mantle.

ISOTOPIC studies of volcanic rocks have greatly enhanced knowledge of the chemistry and evolution of the Earth's mantle. Because of the lesser possibility of crustal contamination, much effort has focused on oceanic volcanics. In general, Sr and Nd isotope ratios in such rocks have proved to be well correlated. The establishment of this linear correlation, or 'mantle array', has been the basis of quantitative models for the present-day chemical structure and evolution of the mantle¹⁻⁴, and for estimates of the planetary ⁸⁷Sr/⁸⁶Sr and Rb/Sr ratios. New data reported here reveal that basalts from some islands have Sr and Nd isotope ratios which deviate significantly from the mantle array. These data thus raise doubts about the validity of some of these models and the estimated Rb/Sr ratio of the Earth. On the other hand, the new data generally support mantle evolution models involving significant transfer of crustal material into the mantle.

Samples from the East Pacific Rise, Mid-Atlantic Ridge, Kerguelen, Samoan Islands, Society Islands, Hawaiian Islands, Easter Island, Sala y Gomez, Guadalupe, Tristan de Cunha, Gough, St Helena and the Azores (Fig. 1) were analysed for isotopic composition of Sr and Nd at the Department of Terrestrial Magnetism, Washington; US Geological Survey, Denver; and Max-Planck-Institut (MPI) für Chemie, Mainz. Conventional mass spectrometric and chemical techniques were used except for some strontium analyses performed at MPI, which were analysed using a double collector system installed in a Finnigan-MAT 261 mass spectrometer. Replicate analysis of samples and standards confirmed that there was no systematic difference between values obtained using single collector and those obtained using double collector. The double collector mode does, however, produce better run statistics in a shorter analysis time. Some samples used in this study have previously been analysed for ⁸⁷Sr/⁸⁶Sr but have been reanalysed because of the higher analytical precision presently obtainable.

Most of the basalts analysed in this study are young—less than 1 Myr old—and hence the measured ratio is the initial ratio. The exceptions are basalts from Kerguelen and the Societies. No age corrections have been applied to these data as exact ages are not known in all cases. The Kerguelen samples range in age up to 27 Myr (ref. 5). In these cases, the measured ⁸⁷Sr/⁸⁶Sr ratios exceed the initial ratios by small, but significant, amounts (up to 0.00010). This difference does not affect the interpretation of the data. In the case of the Society Islands samples, which range in age up to 4 Myr (ref. 6), age corrections are not significant (equal to or less than analytical error).

Results are reported in Table 1 and Fig. 2. We first discuss the data in the context of individual localities before considering their more general implications.

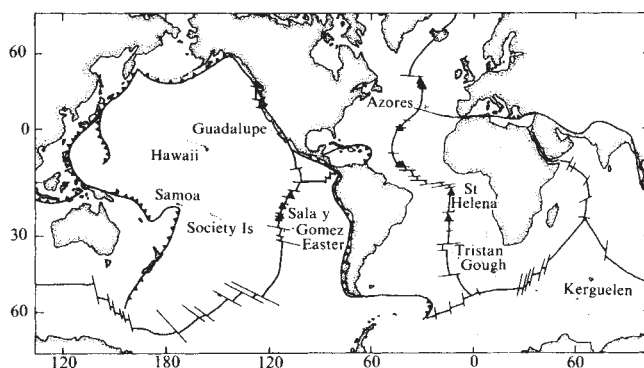


Fig. 1 Plate tectonic map of the world showing locations of dredged basalts (▲) and islands included in this study.

Kerguelen

This island was considered possibly to be a continental fragment but its oceanic nature is now firmly established^{5,7,8}. Dosso *et al.*^{7,8} have found that basaltic rocks from Kerguelen have ¹⁴³Nd/¹⁴⁴Nd ratios, implying source histories ranging from slight light-rare-earth depletion to significant light-rare-earth enrichment (slightly positive to significant negative ϵ_{Nd}). Samples analysed in this study come principally from Foch Island and Corbet Peninsular. Seven of these fall within the range of values reported by Dosso *et al.*, but three, all from Foch Island, have significantly lower ⁸⁷Sr/⁸⁶Sr and higher ¹⁴³Nd/¹⁴⁴Nd ratios than those reported by Dosso. Kerguelen thus encompasses an enormous range of Sr and Nd isotope ratios and includes basalts with Nd isotope ratios both significantly higher and lower than bulk earth. This implies notable mantle heterogeneity on the scale of tens of kilometres. That Foch Island basalt KG5-3 has Sr and Nd isotope ratios similar to basalts from the remainder of Kerguelen but radically different from the flows immediately overlying and underlying it⁵ is further evidence of small-scale heterogeneity.

Hawaii

Sr and Nd isotopic data reported here fall entirely within the range of values reported by O'Nions *et al.*⁹. Our data, and those of O'Nions *et al.* do not reveal any systematic differences in isotopic composition between tholeiitic and alkalic lavas or systematic inter- or intra-island variations, that is while ⁸⁷Sr/⁸⁶Sr and ¹⁴³Nd/¹⁴⁴Nd ratios are quite variable in Hawaii, these variations appear to be largely random. This may, however, be due to the limited amount of data. Two apparent exceptions are (1) basalts from Molokai have notably higher ⁸⁷Sr/⁸⁶Sr than other islands and (2) the Honolulu series volcanics of Oahu have less radiogenic Sr than the Waianae series

§ Present address: Max-Planck-Institut für Chemie, Postfach 3060, 6500 Mainz, FRG.

Table 1 All data reported relative to values of $^{87}\text{Sr}/^{86}\text{Sr} = 0.70800^*$ for the E&A standard and $^{143}\text{Nd}/^{144}\text{Nd} = 0.51265^*$ for the BCR-1 standard

		$^{87}\text{Sr}/^{86}\text{Sr}$	$^{143}\text{Nd}/^{144}\text{Nd}$		$^{87}\text{Sr}/^{86}\text{Sr}$	$^{143}\text{Nd}/^{144}\text{Nd}$
Mid-Atlantic Ridge				Samoa Islands		
HU-197	45°39' N 27°52' W	0.70288 ± 6	0.513109 ± 35	Manu'a Islands		
CH43 106-16 ^{39,13}	45°11' N 27°54' W	0.70338 ± 5	0.513101 ± 20	S17	0.70469 ± 3	0.512785 ± 8
CH43 106-17 ³⁹	45°11' N 27°54' W	0.70307 ± 4	0.513049 ± 30	S221a	{ 0.70444 ± 3	0.512837 ± 7
CH43 106-18 ^{39,13}	45°11' N 27°54' W	0.70337 ± 4	0.513086 ± 20		{ 0.70441 ± 3	0.512831 ± 15
	(leached)	0.70319 ± 5		Tutuila		
CH43 104-7	45°11' N 27°54' W	0.70326 ± 4	0.513054 ± 40	TUT-9 ²⁰	0.70500 ± 2	0.512805 ± 18
CH43 104-59 ³⁹	45°11' N 27°54' W	0.70314 ± 5	0.513096 ± 35	TUT-6D ²⁰ (unleached)	0.70518 ± 3	
HU-56-2	45°13' N 28°00' W	0.70308 ± 4	0.513087 ± 20	(leached)	0.70517 ± 2	0.512860 ± 21
CH13 PD-11 ¹³	50°44' N 29°52' W	0.70264 ± 6	0.513216 ± 25	TL-13 ²⁰	0.70505 ± 1	0.512782 ± 21
A150 RD-20 ^{39,13}	30°01' N 39°04' W	0.70263 ± 6	0.513235 ± 20	ST-26 ²⁰	0.70475 ± 3	0.512816 ± 18
AD2-1 ^{40,41}	20°40' S 13°16' W	0.70285 ± 4	0.513157 ± 19	Upolu		
AD3-3 ^{40,41}	5°47' S 11°25' W	0.70235 ± 2	0.513291 ± 11	UF-1 ¹⁹	{ 0.70490 ± 3	0.512857 ± 21
AD5-5 ^{40,41}	9°39' N 40°27' W	0.70265 ± 4	0.513141 ± 12		{ 0.70493 ± 3	0.512871 ± 17
AD5-11 ^{40,41}	9°39' N 40°27' W	0.70284 ± 3	0.513163 ± 11	UF-2 ¹⁹	{ 0.70500 ± 4	0.512935 ± 9
					{ 0.70583 ± 1	0.512699 ± 13
East Pacific Rise				UP-1 ¹⁹ (unleached)	{ 0.70584 ± 1	0.512705 ± 21
Amph D-4 ¹⁵	18°52' S 113°19' W	0.70257 ± 2	0.513107 ± 10	(leached)	{ 0.70582 ± 2	0.512702 ± 8
PD1-1 ^{41,42}	7°47' S 108°10' W	0.70251 ± 5	0.513211 ± 17	US-1 ¹⁹	{ 0.70583 ± 3	0.512746 ± 14
PD3A ^{41,42}	12°52' S 110°57' W	0.70252 ± 2	0.513202 ± 20	UM-2 ¹⁹	{ 0.70562 ± 2	0.512678 ± 17
PD4G ^{41,42}	18°25' S 113°20' W	0.70253 ± 2	0.513163 ± 11		{ 0.70562 ± 2	0.512669 ± 27
				UPO-2 ²⁰ (unleached)	{ 0.70486 ± 2	0.512784 ± 14
Hawaiian Islands				(leached)	{ 0.70483 ± 2	0.512821 ± 15
Loihi Seamount				UPO-9 ²⁰	{ 0.70547 ± 2	0.512750 ± 8
Dredge 2		0.70350 ± 2	0.512979 ± 14	UPO-F-9-1 ²⁰ (unleached)	{ 0.70494 ± 2	0.512905 ± 20
Hawaii				(leached)	{ 0.70491 ± 2	0.512922 ± 20
BHVO-1 Kilauea 1919		0.70346 ± 2	0.512948 ± 13	UPO-F-9-8 ²⁰	{ 0.70540 ± 1	0.512796 ± 16
		0.70345 ± 5	0.512990 ± 24	JKU-1 ²⁰	{ 0.70552 ± 1	0.512615 ± 24
KH-11 ¹¹ Kohala		0.70354 ± 2	0.512979 ± 32		{ 0.70565 ± 1	0.512655 ± 10
		0.70357 ± 2	0.512988 ± 32	JKU-21 ²⁰	{ 0.70559 ± 1	0.512642 ± 15
MG-1a ¹¹ Kohala		0.70361 ± 2	0.512973 ± 30	Savai'i		
W8 ¹¹ Kohala		0.70358 ± 2	0.512973 ± 13	VA-1 ¹⁹ (unleached)	{ 0.70635 ± 2	0.512775 ± 18
Molokai				(leached)	{ 0.70636 ± 2	0.512763 ± 17
GA 572 ⁴³		0.70382 ± 4	0.512944 ± 10		{ 0.70635 ± 2	0.512761 ± 8
WMOL-1 ²²		0.70411 ± 2†	0.512943 ± 18	VM-2 ¹⁹	{ 0.70651 ± 2	0.512734 ± 16
Oahu (Waianae Series)				VP-1 ¹⁹	{ 0.70601 ± 3	0.512689 ± 12
5X-667-3 ²²		0.70378 ± 2†	0.513036 ± 24		{ 0.70602 ± 2	0.512694 ± 16
5X-197-2 ²²		0.70380 ± 2	0.512982 ± 10	VP-2 ¹⁹ (unleached)	{ 0.70575 ± 2	0.512750 ± 14
GA553 ⁴³		0.70381 ± 3	0.512935 ± 11	(leached)	{ 0.70576 ± 2	0.512777 ± 18
GA557 ⁴³		0.70363 ± 6	0.512986 ± 9		{ 0.70577 ± 2	0.512757 ± 22
Kauai				SAV-1 ²⁰	{ 0.70573 ± 3	0.512729 ± 20
GA565 ⁴³		0.70371 ± 4	0.512973 ± 9	Society Islands		
GA649 ⁴³		0.70359 ± 4	0.513019 ± 28	Tahiti		
ID-872-2 ²²		0.70369 ± 3†	0.513026 ± 34	74-416 ¹⁸	0.70434 ± 2	0.512850 ± 8
Nihoa				74-422 ¹⁸ (unleached)	0.70419 ± 2	0.512931 ± 20
8G141-2 ²²		0.70364 ± 3†	0.513068 ± 20	(leached)	0.70418 ± 2	0.512918 ± 18
La Perouse Pinnacles				Hauhine		
LPP-W-30 ²²		0.70371 ± 4	0.513014 ± 9	73-95 ¹⁸ (unleached)	0.70567 ± 2	0.512704 ± 14
				(leached)	0.70563 ± 1	0.512674 ± 8
Kerguelen				Bora Bora		
Foch Island				73-332 ¹⁸	0.70495 ± 1	0.512829 ± 17
KG2-1 ⁵		0.70388 ± 5	0.512878 ± 30	Maupiti		
KG4-1 ⁵		0.70407 ± 6	0.512870 ± 30	73-204 ¹⁸	0.70530 ± 1	0.512706 ± 10
KG5-3 ⁵		{ 0.70532 ± 2	0.512631 ± 24	Tahaa		
		{ 0.70530 ± 2	0.512641 ± 12	73-185 ¹⁸ (unleached)	0.70691 ± 2	0.512587 ± 26
KG6-2 ⁵		0.70398 ± 4	0.512907 ± 25	(leached)	0.70694 ± 2	0.512571 ± 16
Rest of Island				73-186 ¹⁸ (unleached)	0.70504 ± 2	0.512804 ± 36
KG8-4 ⁴ (unleached)		0.70531 ± 4	0.512645 ± 40	(leached)	0.70505 ± 2	0.512803 ± 20
(leached)		0.70534 ± 6		Gough Island		
KG11-1 ⁵		0.70526 ± 6	0.512671 ± 30	G-8 ¹⁵	0.70532 ± 6	0.512574 ± 20
KG14-3 ⁵		0.70532 ± 5	0.512631 ± 35	Tristan da Cunha		
I75		0.70538 ± 4	0.512685 ± 20	Tr-1 ^{15,45} (NMNH 110014)	0.70505 ± 2	0.512534 ± 10
I380		0.70531 ± 6	0.512657 ± 25	Tr-4 ^{15,45} (NMNH 110017)	0.70509 ± 2	0.512526 ± 15
I381		0.70521 ± 6	0.512671 ± 15	Tr-6 ^{15,45} (NMNH 110019)	0.70501 ± 2	0.512567 ± 10
Easter Island				St Helena		
17732 ⁴⁴ (Poike)		0.70326 ± 5	0.512992 ± 30	2928 ^{15,46}	0.70287 ± 3	0.512883 ± 18
EC-307 ⁴⁴ (Poike)		0.70319 ± 6	0.513062 ± 25	2876 ⁴⁶ (NMNH 99659)	0.70287 ± 2	0.512853 ± 29
EH-39 ⁴⁴ (Poike)		0.70328 ± 4	0.513056 ± 25	2893 ⁴⁶ (NMNH 99661)	0.70292 ± 2	0.512843 ± 32
EH-20 ⁴⁴ (Rano Kau)		0.70321 ± 5	0.513004 ± 30	2926 ⁴⁶ (NMNH 99653)	0.70285 ± 2	0.512871 ± 32
KK72-43-7 ⁴⁴		0.70324 ± 4	0.513012 ± 35	2882 ⁴⁶ (NMNH 99658)	0.70292 ± 1	0.512857 ± 12
EC-385 ⁴⁴ (Terevaka)		0.70301 ± 5	0.513008 ± 30	102 (NMNH 109984)	0.70296 ± 2	0.512842 ± 14
Sala y Gomez				Azores		
Y73-4-30-20 ⁴⁴		0.70327 ± 5	0.512953 ± 20	Sao Miguel		
Guadalupe				SM-2A ¹⁶	0.70337 ± 6	0.512926 ± 30
GU77 ¹⁵		0.70326 ± 4	0.512942 ± 25	SM-12 ¹⁶	{ 0.70514 ± 1	0.512707 ± 11
				SM-6 ¹⁶	{ 0.70430 ± 2	0.512812 ± 10
					{ 0.70429 ± 2	0.512810 ± 18
				SM-28 ¹⁶	{ 0.70525 ± 6	0.512732 ± 28
				Faial		
				F-2 ¹⁶	0.70386 ± 2	0.512871 ± 24
				F-14 ¹⁶	0.70394 ± 6	0.512870 ± 28
				F-21 ¹⁶	0.70384 ± 6	0.512943 ± 25
				F-29 ¹⁶	0.70392 ± 2	0.512877 ± 18
				F-33 ¹⁶	0.70393 ± 6	0.512843 ± 24

* Measured values at the Max-Planck-Institut. These values were 0.70792 and 0.51271, and 0.70808 and 0.51273 at the Department of Terrestrial Magnetism, Washington and the US Geological Survey, Denver respectively. Error quoted are 2 s.e.m. based on in-run statistics. This statistic probably slightly overestimates reproducibility, which is better evaluated from the replicate analyses. Superscripts refer to other published information of the sample.

† Sr data from ref. 22.

of that island. These distinctions do not apply to Nd. On Hawaii, the individual volcanoes are not isotopically distinct, although the Kohala lavas are notably isotopically uniform, in agreement with Feigenson *et al.*¹⁰. The easternmost and youngest volcano of the chain, Loihi, has Sr and Nd isotope ratios typical of the chain in general.

East Pacific Rise and Islands

The East Pacific Rise, Easter Island, Sala y Gomez and Guadalupe all have similar isotopic compositions which fall within the range of values from the Galapagos¹¹. The Poike and Ranu Kau basalts from Easter Island exhibit variations only slightly greater than analytical error, while EC-385 of the Terevaka volcanics has a significantly lower $^{87}\text{Sr}/^{86}\text{Sr}$ ratio. Previously reported $^{87}\text{Sr}/^{86}\text{Sr}$ ratios from Easter (summarized in ref. 12) show much more dispersion, perhaps due to the relatively large analytical error of the earlier data. Samples from the East Pacific Rise have isotopic compositions similar to other mid-ocean ridge basalts (MORB).

Mid-Atlantic Ridge and South Atlantic Islands

Except for those samples from the vicinity of 45°N, the Mid-Atlantic Ridge samples have typical MORB compositions. The 45°N area was known to be anomalous from previous studies¹³. The relatively high $^{87}\text{Sr}/^{86}\text{Sr}$ and lower $^{143}\text{Nd}/^{144}\text{Nd}$ ratios of the 45°N samples are thus not surprising. Note that the $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of CH43 106–18 were lower after HCl leaching, indicative of seawater contamination. This may also account for the high $^{87}\text{Sr}/^{86}\text{Sr}$ ratio of CH43 106–16. Three new analyses from Tristan da Cunha fall within the range of values previously reported^{9,14}, and although there are still relatively few data from this island, it appears that the volcanics are isotopically homogeneous. Gough Island appears to be isotopically similar to but distinct from Tristan. Both Tristan and Gough have $^{143}\text{Nd}/^{144}\text{Nd}$ ratios lower than chondritic, implying time-integrated light-rare-earth enrichment of their sources. The six St Helena samples analysed in this study suggest this island also has an isotopically uniform source but one which is quite distinct from Tristan and Gough (A. Zindler, personal communication, however, reports samples with distinctly higher $^{87}\text{Sr}/^{86}\text{Sr}$). The St Helena data fall off the mantle array and are unique in that they plot to the left of it. St Helena is also unique in having extremely high $^{206}\text{Pb}/^{204}\text{Pb}$ ratios¹⁵.

Azores

Most Azores islands basalts have $^{87}\text{Sr}/^{86}\text{Sr}$ ratios between 0.70332 and 0.70354 but basalts from Sao Miguel can have ratios up to 0.70525 (ref. 16). Hawkesworth *et al.*¹⁷ found that $^{87}\text{Sr}/^{86}\text{Sr}$ and $^{143}\text{Nd}/^{144}\text{Nd}$ ratios from Sao Miguel define a trend which is distinct from the mantle array. New data reported here confirm the results of Hawkesworth *et al.* Faial, which after Sao Miguel has the highest $^{87}\text{Sr}/^{86}\text{Sr}$ in the Azores, has isotopic compositions which plot at the intersection of the Sao Miguel trend and the mantle array.

Society Islands

Duncan *et al.*^{6,18} found that these islands become progressively older to the west, ranging in age from 1.0 to 4.3 Myr and that their Sr is highly radiogenic with $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of 0.7036–0.7067. $^{87}\text{Sr}/^{86}\text{Sr}$ and $^{143}\text{Nd}/^{144}\text{Nd}$ ratios fall on the trend of the Sao Miguel data but extend it to much higher $^{87}\text{Sr}/^{86}\text{Sr}$ and lower $^{143}\text{Nd}/^{144}\text{Nd}$ ratios. Sample 73–185 from Tahaa, at one extreme of this trend, deviates markedly from the mantle array. There is no apparent systematic variation either in time or in space, the island of Tahaa exhibiting more variation than the other islands combined.

Samoa

Like basalts from Sao Miguel and the Society Islands, many Samoan basalts have $^{87}\text{Sr}/^{86}\text{Sr}$ and $^{143}\text{Nd}/^{144}\text{Nd}$ ratios which

plot off the mantle array. However, unlike Society and Sao Miguel basalts, Samoan basalts do not form a linear trend but scatter about the Sao Miguel–Society trend. Hedge *et al.*¹⁹ found that $^{87}\text{Sr}/^{86}\text{Sr}$ ratios increased westwards along the island chain. This is borne out by the new data, which include many of the samples analysed by Hedge *et al.* The geographical variation is not strong, however, as, for example, Tutuila basalts plot entirely within the Upolu field.

Samoa volcanism has occurred in two phases^{20,21}: the earlier, or pre-erosional volcanism is normal island-chain type, that is becoming progressively younger to the east, but a younger episode of post-erosional volcanism has occurred along a fissure connecting the three westernmost and oldest islands. Both pre- and post-erosional volcanics have Sr and Nd isotope ratios which plot off the mantle array, and there is no apparent difference in the extent to which the two types deviate from the mantle array, although the post-erosional lavas generally have the highest $^{87}\text{Sr}/^{86}\text{Sr}$ and lowest $^{143}\text{Nd}/^{144}\text{Nd}$ ratios. Basalts from the Manua islands, at the eastern end of the chain, plot within the mantle array.

Contamination during or after magma ascent

Possible causes for the divergence of isotope ratios of some of the islands from the mantle array include post-eruptional weathering effects and contamination by altered basalt or sediment of the oceanic crust during magma ascent. Acid-leaching, which is useful in identifying samples affected by post-eruptional alteration (see, for example, ref. 8), reveals no difference between leached and unleached fractions (Table 1). This argues persuasively against the former possibility. Arguments against the latter have been given by Duncan and Compston¹⁸ and by Hawkesworth *et al.*¹⁷. To these may be added the following observation. The Samoan basalts tend to be primitive in nature (that is, high Mg and Ni)¹⁹, which suggests they did not spend any considerable time in a crustal magma chamber and hence are particularly unlikely to have suffered crustal contamination. Although oceanic crustal contamination during magma ascent cannot be ruled out entirely, it is unlikely to be the cause of the observed isotopic variations, which are considered to reflect chemical and isotopic heterogeneity in the mantle.

Intra-island and intra-archipelago variations

Figure 2 illustrates the variability of isotopic heterogeneity of oceanic islands. Single isolated islands such as St Helena, Tristan da Cunha, Ascension and Bouvet (Table 1 and ref. 9) tend to

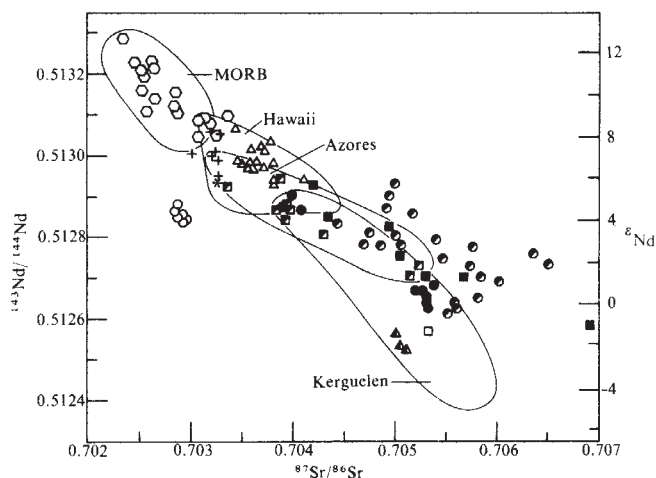


Fig. 2 Results obtained in this study plotted on a $^{87}\text{Sr}/^{86}\text{Sr}$ versus $^{143}\text{Nd}/^{144}\text{Nd}$ variation diagram. Fields for mid-ocean-ridge basalts (MORB), Azores, Hawaii and Kerguelen are drawn on the basis of these and previously published data. ○, MORB; +, Easter Island and Sala y Gomez; *, Guadalupe; △, Hawaii; ○, St Helena; ▣, Azores; ●, Samoa; ■, Societies; □, Gough; ▲, Tristan da Cunha; and ●, Kerguelen. Much of the data from Samoa fall outside the boundary of the conventional mantle array and data from Sao Miguel (Azores) and the Societies define a trend at low angle to the mantle array.

be isotopically homogeneous while islands or island groups associated with large oceanic plateaus, such as Iceland, Kerguelen, Hawaii, the Galapagos and the Azores (Table 1 and refs 8, 9, 11, 17, 22, 23), exhibit more variation, as do the Pacific island chains, such as Hawaii, Samoa and the Societies.

However, the apparent homogeneity of some islands may be related to sampling problems, and the question of local isotopic variations on oceanic islands requires a much more systematic approach to sampling than has previously been undertaken.

These intra-island and intra-archipelago variations may result partially or largely from mixing of an enriched source (or a melt derived from such a source) with depleted mantle. The apparent homogeneity of some islands indicates that such mixing is not ubiquitous. Furthermore, in the case of Samoa, Sr–Nd and particularly Sr–Pb relationships²⁴ suggest heterogeneity of the source is due at least in part to factors other than mixing with depleted mantle.

Mantle evolution

Explanations for the Nd–Sr relationship which have been suggested include mixing between a primitive or planetary reservoir having $^{87}\text{Sr}/^{86}\text{Sr}$ of ~ 0.705 and $^{143}\text{Nd}/^{144}\text{Nd}$ of ~ 0.51265 and a depleted reservoir with the isotopic characteristics of MORB (that is $^{87}\text{Sr}/^{86}\text{Sr} = 0.7027$ and $^{143}\text{Nd}/^{144}\text{Nd} = 0.51315$) and variable depletion of an initially homogeneous planetary reservoir, the most severely depleted portion represented by the MORB reservoir²⁻⁴. A basic assumption of these models is that the relationship between Sr and Nd isotopic ratios in oceanic basalts is monotonic. That the $^{87}\text{Sr}/^{86}\text{Sr}$ ratio of the bulk earth may be found by taking the intersection of the 'mantle array' with the chondritic $^{143}\text{Nd}/^{144}\text{Nd}$ is implicit.

The new data presented here and those of Hawkesworth *et al.*¹⁷ clearly violate the assumption of a single linear relationship describing Nd and Sr isotopic variations in oceanic basalts. We thus have to choose whether to retain the established assumptions and explain the new data as deviations from the rule, or to suspend the assumption of the linear nature of the data and

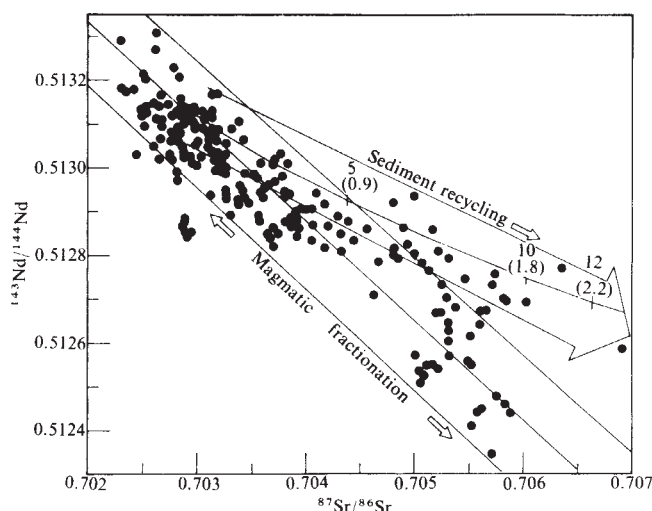


Fig. 3 Variation of $^{143}\text{Nd}/^{144}\text{Nd}$ with $^{87}\text{Sr}/^{86}\text{Sr}$ in oceanic basalts. The data clearly define a wedge-shape rather than linear array. One possible explanation is illustrated: magmatic fractionation of Rb/Sr and Sm/Nd (perhaps principally that occurring at mid-ocean-ridges, as suggested in refs 33–35) in time results in a linear band of data while injection of continent-derived sediment into the mantle results in deviation from this linear pattern. Also shown is the effect of mixing a hypothetical ancient oceanic sediment (Rb/Sr and Sm/Nd identical to the deep-sea red clay analysed by McCulloch and Wasserburg³⁷ but having a model age of 2,000 Myr, that is present $^{87}\text{Sr}/^{86}\text{Sr} = 0.7239$ and $^{143}\text{Nd}/^{144}\text{Nd} = 0.51204$) with a reservoir having the isotopic composition of MORB. This sediment would be roughly appropriate with a 1,800 Myr recycling time for example, refs 35, 36. Ticks and numbers along the mixing line refer to the per cent sediment in the mixture. Numbers in parentheses refer to the case where the depleted end member is mantle with 20 p.p.m. Sr, those without parentheses to mixing with basalt having 120 p.p.m. Sr. Sources of data: Table 1, refs 8, 9, 17, 11, 23, 47–51 and W.M.W. (unpublished).

accept the new data into the mantle array without special pleading. If we take the latter course, the mantle array becomes a rather broad band of data where several subsets (from individual island groups) have trends with substantially different slopes. In that case, the assumption of a basically binary system cannot be maintained, and we cannot make any quantitative estimates about a point representing bulk earth. For example, it is quite reasonable to think that the Earth has differentiated to produce six geochemically distinct reservoirs: (1) an upper continental crust; (2) a lower continental crust, (3) an oceanic crust; (4) a depleted mantle region; (5) an undepleted mantle region; and (6) the core. The conventional linear mantle array might then be produced by mixing of (4) and (5), (3) and (4), or perhaps even (1) and (4). The different trends within the mantle array may then be the result of different combinations, but there is no requirement that the bulk earth point must lie within the array. Similar arguments can be made for Pb isotopes, which, on the $^{207}\text{Pb}/^{204}\text{Pb}$ versus $^{206}\text{Pb}/^{204}\text{Pb}$ diagram, show an overall positive correlation. These can be resolved into several different individual subparallel trends for individual island groups, but virtually all the data lie outside the locus of possible values of bulk earth Pb isotope ratios.

The data from Sao Miguel, Samoa and the Societies are offset from the conventional mantle array in the direction of continental crust and in this sense, the involvement of some kind of crustal component is implied, as Hawesworth *et al.*¹⁷ suggested with reference to Sao Miguel. If this crustal component is not added during ascent of the magma, then it must be present in the source. The possibility of transport of crustal material into the mantle, suggested by Armstrong²⁵, has recently received wide consideration (see refs 1, 26–36), and Hofmann and White^{34,35} and Chase³⁶ have specifically proposed that recycled oceanic crust is the source of oceanic island basalts. An interpretation of the Sr and Nd isotope data consistent with recycling of oceanic crust is shown in Fig. 3. Magmatic processes operating during the formation of oceanic lithosphere produce correlated variations in Rb/Sr and Sm/Nd. Most of this lithosphere is ultimately returned to the deeper mantle and in time, correlated variations in $^{87}\text{Sr}/^{86}\text{Sr}$ and $^{143}\text{Nd}/^{144}\text{Nd}$ result which define the conventional linear mantle array^{33–35}. Addition (and subtraction) of continent derived sediment or seawater interaction with crust will, however, produce deviations to the right of the mantle array. The effect of mixing between a hypothetical 'ancient' oceanic sediment and a MORB-type composition is illustrated. As Fig. 3 shows, such mixing can at least qualitatively account for those sources with compositions outside the conventional mantle array and that relatively small amounts of sediment are required. In this interpretation, whether a particular unit of recycled crust has compositions lying outside or inside the conventional mantle array depends on the extent to which the crust has interacted with seawater and/or the extent to which associated sediments are subducted. Sources lying inside the conventional mantle array, such as Tristan and Kerguelen, may represent recycled crust with little or no subducted sediment.

The mixing curve shown in Fig. 3 is illustrative only. The range of isotopic compositions in modern oceanic sediments, much less ancient ones, is not well known; thus any attempt at more quantitative modelling is not justified. It is clear, however, that sediments can be highly variable^{37,38}, both in isotopic composition and in Rb/Sr, Sm/Nd and Sr/Nd ratios (the latter determining the curvature of mixing curves). Even compositions lying inside the conventional mantle array, for example Tristan and Kerguelen, could be explained by mixing between a sediment, for example the North American Shale Composite³⁷, and depleted mantle.

Seawater is also a possible 'continental contaminant', as mentioned above. Sediment has been emphasized here because of the correlated variations in Sr and Nd isotope ratios in the Society Islands and Sao Miguel trends seem to suggest a contaminant containing both Sr and Nd, whereas seawater, which contains essentially no Nd, would be expected to affect only

the Sr isotope ratios. On the other hand, the correlation between Sr and Nd isotope ratios in the Samoan source is weak, which might indicate that seawater is the contaminant.

It is also possible that reservoirs having isotope ratios lying outside the mantle array result from purely magmatic processes. Diagrams illustrating the effects of partial melting on Rb/Sr and Sm/Nd ratios (for example Fig. 5 of ref. 9; Fig. 4 of ref. 17) reveal that these sources could be produced as partial melts of a previously depleted mantle. This possibility is also consistent with recycling of oceanic crust as partial melting of a previously depleted reservoir occurs during MORB genesis.

Models involving recycling of oceanic crust thus provide several possible means of explaining both those data falling outside the conventional mantle array and those sources having $^{143}\text{Nd}/^{144}\text{Nd}$ ratios lower than bulk earth. Neither of these features are consistent with binary mixing between a depleted and an undepleted reservoir, such as advocated by Wasserburg and DePaolo³, or with variable depletion of a primordial reservoir⁴.

Conclusions

In view of the above new data, as well as those of Hawkesworth *et al.*¹⁷, recycling of oceanic crust into the mantle deserves serious consideration as an important process in the chemical

evolution of the Earth. However, Pb, Hf, He, Ar as well as Sr and Nd isotopic data, chemical data and geophysical and chemical plausibility must be carefully considered before such models are accepted. The present data support such models rather than confirm them. On the other hand, these new data cannot be explained in terms of simple mixing between an undepleted planetary mantle reservoir and a depleted one, nor can Pb isotope data. Such an undepleted reservoir, though it may exist, must have only a very minor role in present crust-mantle interaction. The data are also not explained solely by variable magmatic depletion of an initially homogeneous mantle.

Finally, until the complexities of Sr-Nd isotope systematics in the mantle are better understood, values of $^{87}\text{Sr}/^{86}\text{Sr}$ and Rb/Sr for the bulk earth estimated from Sr-Nd relationships in oceanic basalts must be viewed with scepticism.

We thank J. Clark, J. Dymond, R. Duncan, M. Feigenson, M. Lanphere, A. McBirney, W. Melson and H. Banks of the Smithsonian Institution, J. Natland, Z. Peterman, C. Hedge, J.-G. Schilling, H. Sigurdsson and S. Sun for providing samples. This work benefited from discussions with, critical reviews by and assistance of M. Tatsumoto, H. Feldmann, C. Göpel, R. Gross, E. Ito, E. Jagoutz, D. James, J. Patchett, Z. Peterman, S. Sun and H. Voshage.

Received 8 February; accepted 15 March 1982.

- O'Nions, R. K., Evensen, N. M. & Hamilton, P. J. *J. geophys. Res.* **84**, 6091-6101 (1979).
- Allègre, C. J., Ben Othman, D., Polve, M. & Richard, P. *Earth planet. Sci. Lett.* **19**, 293-306 (1979).
- Wasserburg, G. J. & DePaolo, D. J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3594-3598 (1979).
- Jacobsen, S. B. & Wasserburg, G. J. *J. geophys. Res.* **84**, 7411-7427 (1979).
- Watkins, N. D., Gunn, B. M., Nougier, J. & Baksi, A. K. *Bull. geol. Soc. Am.* **85**, 201-212 (1974).
- Duncan, R. A. thesis, Australian National Univ. (1975).
- Dosso, L. *et al. Earth planet. Sci. Lett.* **43**, 46-60 (1979).
- Dosso, L. & Murthy, V. R. *Earth planet. Sci. Lett.* **48**, 268-276 (1980).
- O'Nions, R. K., Hamilton, P. J. & Evensen, N. M. *Earth planet. Sci. Lett.* **34**, 13-22 (1977).
- Feigenson, M. D., Hofmann, A. W. & Spera, F. J. (in preparation).
- White, W. M. & Hofmann, A. W. *Yb. Carnegie Instn Wash.* **77**, 596-606 (1978).
- Bonnati, E., *et al. J. geophys. Res.* **82**, 2457-2478 (1977).
- White, W. M. & Schilling, J.-G. *Geochim. cosmochim. Acta* **42**, 1501-1516 (1978).
- O'Nions, R. K. & Pankhurst, R. J. *J. Petrol.* **15**, 603-634 (1974).
- Sun, S.-S. *Phil. Trans. R. Soc. A297*, 409-445 (1980).
- White, W. M., Tapia, M. D. M. & Schilling, J.-G. *Contr. Miner. Petrol.* **69**, 201-213 (1979).
- Hawkesworth, C. J., Norry, M. J., Roddick, J. C. & Vollmer, R. *Nature* **280**, 28-31 (1979).
- Duncan, R. A. & Compston, W. *Geology* **4**, 728-732 (1976).
- Hedge, C. E., Peterman, Z. E. & Dickinson, W. R. *Bull. geol. Soc. Am.* **83**, 2709-2714 (1972).
- Natland, J. H. thesis, Univ. of California, San Diego (1975).
- Natland, J. H. *Am. J. Sci.* **280A**, 709-735 (1980).
- Lanphere, M. A., Dalrymple, G. B. & Clague, D. A. *Init. Rep. DSDP Leg 40*, 695-706 (1980).
- Zindler, A., Hart, S. R., Frey, F. A. & Jakobsson, S. P. *Earth planet. Sci. Lett.* **45**, 249-262 (1979).
- Todt, W. & White, W. M. *Terra Cognita* **1** (in the press).
- Armstrong, R. L. *Rev. Geophys.* **6**, 175-199 (1968).
- Doe, B. R. & Zartman, R. E. in *Geochemistry of Hydrothermal Ore Deposits*, 2nd edn (ed. Barnes, H. L.) 22-70 (1979).
- Richter, F. M. & Ribe, N. M. *Earth planet. Sci. Lett.* **43**, 212-222 (1979).
- DePaolo, D. J. & Wasserburg, G. J. *Geochim. cosmochim. Acta* **43**, 615-627 (1979).
- Allègre, C. J., Brévart, O., Dupré, B. & Minster, J. F. *Phil. Trans. R. Soc. A297*, 447-477 (1980).
- Davies, G. F. *Nature* **290**, 208-213 (1981).
- Armstrong, R. L. *Phil. Trans. R. Soc. A301*, 443-472 (1981).
- O'Nions, R. K. & Hamilton, P. J. *Phil. Trans. R. Soc. A301*, 473-487 (1980).
- Dosso, L. & Murthy, V. R. *Earth planet. Sci. Lett.* (submitted).
- Hofmann, A. W. & White, W. M. *Yb. Carnegie Instn Wash.* **79**, 477-483 (1980).
- Hofmann, A. W. & White, W. M. *Earth planet. Sci. Lett.* **57**, 421-436 (1982).
- Chase, C. G. *Earth planet. Sci. Lett.* **52**, 277-284 (1981).
- McCulloch, M. T. & Wasserburg, G. J. *Science* **200**, 1003-1011 (1978).
- Goldstein, S. L. & O'Nions, R. K. *Nature* **292**, 324-327 (1981).
- Schilling, J.-G. *et al. Am. J. Sci.* (submitted).
- Engel, A. E. J. & Engel, C. G. *Science* **144**, 1330-1333 (1964).
- Tatsumoto, M. *Earth planet. Sci. Lett.* **38**, 63-87 (1978).
- Engel, A. E. J. & Engel, C. G. *Science* **146**, 477-485 (1964).
- Cooper, J. A. & Richards, J. R. *Earth planet. Sci. Lett.* **1**, 259-269 (1966).
- Clark, J. G. & Dymond, J. J. *Volcanol. Geothermal Res. J.* **2**, 29-48 (1977).
- Baker, P. E., Harris, I. G. & LeMaitre, R. W. *Phil. Trans. R. Soc. A256*, 439-575 (1964).
- Daly, R. A. *Proc. Am. Acad. Arts Sci.* **62**, 31-92 (1927).
- Richard, P., Shimizu, N. & Allègre, C. J. *Earth planet. Sci. Lett.* **31**, 269-278 (1976).
- DePaolo, D. J. & Wasserburg, G. J. *Geophys. Res. Lett.* **3**, 743-746 (1976).
- White, W. M. *Yb. Carnegie Instn Wash.* **78**, 325-331 (1979).
- Cohen, R. S., Evensen, N. M., Hamilton, P. J. & O'Nions, R. K. *Nature* **283**, 149-153 (1980).
- White, W. M., *et al. Geol. Soc. Am. Abstr. Progr.* **13**, 580 (1981).

Kinematics of oceanic thrusting and subduction from basal sections of ophiolites

F. Boudier, A. Nicolas & J. L. Bouchez

Laboratoire de Tectonophysique, Université de Nantes, rue de la Houssinière, 44072 Nantes Cedex, France

The structural analysis of high stress deformation at the base of the peridotite section in many ophiolites in suture zones enables the kinematics of oceanic convergence responsible for the suture to be defined. This deformation results from the thrusting over oceanic crust of a young lithosphere involved in a subduction process. Flow plane, flow line, sense of shear related to the thrusting are geometrically deduced from the study of mineral preferred orientations in the basal peridotite and in the quartzites belonging to the amphibolite sole.

In the palimpsestic reconstruction of palaeotectonics, a key problem is to determine kinematics of plates (direction and sense of relative motion) at destructive margins. At present, the following information is available.

First, the direction of the palaeo suture zone is assumed to be parallel to the trend of ophiolite and blueschist lineaments and to the alignment of eruptive bodies which are ascribed to

an active margin magmatism. However, no information is available on the direction of motion of the subducting plate. This direction may be nearly parallel to the suture direction as is the case along the western boundary of the Pacific plate¹ and along the eastern segment of the Hellenic trench system².

Second, the sense of the palaeo subduction is deduced from: (1) the different geological characters between the active and

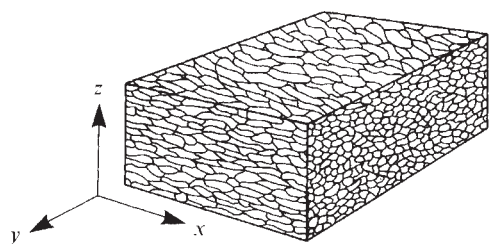


Fig. 1 x, y, z axes of the finite strain defined from the shape fabric: foliation (xy plane) and mineral lineation (x direction).

passive margins of the suture. The active margin is characterized by arc sedimentation, andesitic-granodioritic magmatism, and K/Na polarity, whereas the passive margin is characterized by a sedimentation of continental rise type. Such a simple scheme may be obscured by complex collision situations. (2) The sense of thrusting of the ophiolites which is inferred from their position with regard to the former oceanic crust from which they derived. This criterion is still ambiguous as obduction has been related to either sense of subduction³⁻⁵. Moreover it is not always possible to recognize a thrust whose kinematics is connected to the subduction process from one due to later nappe emplacement, possibly driven by gravity.

A new approach to this problem is now proposed, based on the study of basal tectonites and metamorphic aureoles of many ophiolites. These formations have been ascribed to oceanic thrusts⁶⁻⁸. Kinematic analysis of deformation mainly in the peridotite mylonites makes it possible to deduce the direction and sense of transport of the upthrust lithosphere which may then be related to the direction of convergence of the plates.

Large homogeneous deformations

The kinematic analysis of homogeneous plastic deformation has recently been developed, in particular in peridotites^{9,10} and in quartzites¹¹.

In rocks deformed by homogeneous plastic flow, we accept the concept that the principal axes of the finite strain ellipsoid $x \geq y \geq z$ coincide with the principal axes of the shape anisotropy of minerals, with z perpendicular to the foliation and x parallel to the mineral lineation (Fig. 1). In peridotites, the principal directions of strain are marked by the shape of the olivine crystals, or better, in the field, by spinel or feldspar aggregates. In quartzites they are marked by the shape of quartz crystals, often emphasized by phyllitic bands.

Fabric measurements in peridotites and quartzites show that the mean orientation of slip directions and slip planes in the principal minerals (olivine and quartz respectively) are related to the axes (x, y, z) (Fig. 1) of the finite strain ellipsoid so that the mean direction of slip commonly lies at small angle to the x axis (Fig. 2a). This obliquity has been observed in numerous cases and in various materials; it has been attributed to shear flow (see refs 10, 12), with the inference that the flow plane and the flow line can be equated with the mean orientation of the slip planes and the slip lines respectively in the principal rock-forming minerals (Fig. 2b). The sense of shear flow is deduced from the sense of the obliquity. This interpretation has been confirmed by planar simulation of simple shear¹³, by studies in natural shear zones where the flow geometry is well known¹⁴⁻¹⁶ and recently by experimental deformation in ice¹⁷.

The obliquity of the kinematic directions relative to the x and z axes of the finite strain ellipsoid can be deduced from petrofabric measurements (Fig. 2b), or more simply from thin sections cut parallel to the xz plane observed between crossed polarizers (Fig. 2a).

Oceanic thrusting conditions

The basal peridotites in many ophiolite complexes (see refs 7, 18) overlie a sole of metamorphic rocks with facies grading downwards from granulite and/or amphibolite facies to greenschist facies. This metamorphism is developed in oceanic crust formations as a result of mechanical and thermal effect of an overriding slice of oceanic lithosphere^{6,7,19}. Thirty ophiolite

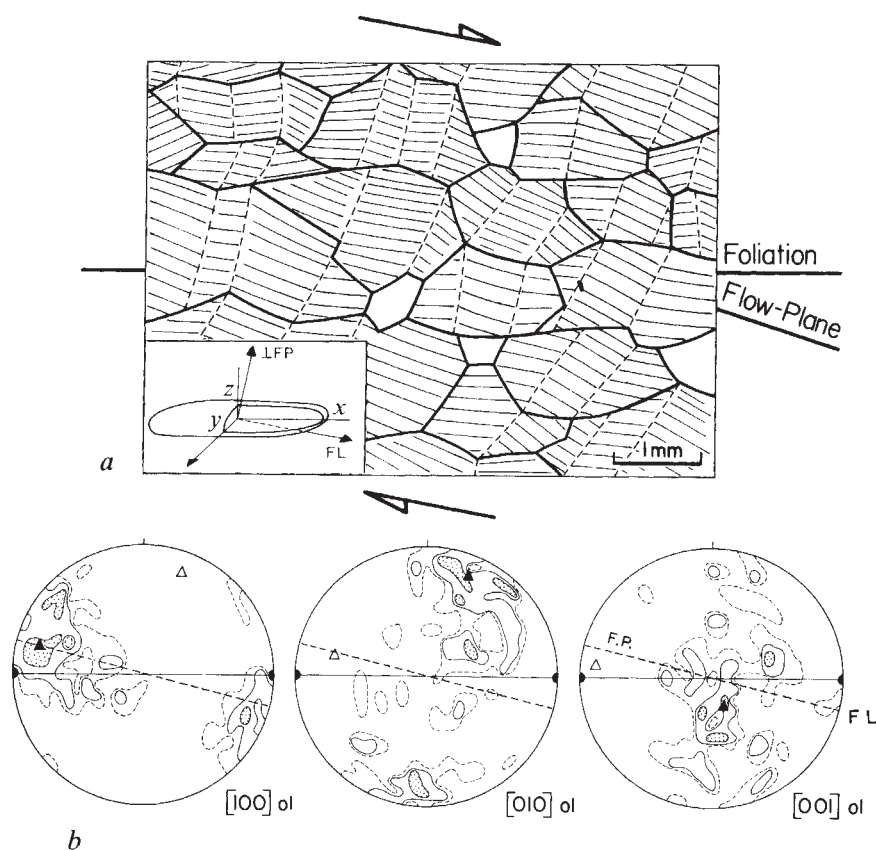


Fig. 2 Schematic section in the xz plane of an olivine or quartz aggregate. Dashed lines are the traces of sub-boundaries (visible under crossed polarizers), perpendicular to the slip direction. Solid lines are the traces of slip planes (not directly visible under crossed polarizers). The sense of shear is deduced from the orientation of foliation relative to average slip plane (equated to the flow plane, FP) and/or lineation relative to average slip direction (equated to the flow line, FL). *b*, Equal area projection (lower hemisphere) of the olivine crystallographic axes. The flow plane (FP) (dashed line) is equated with the average slip plane (010_{ol}), the flow line (FL), with the average slip direction (100_{ol}). Solid line, foliation; ●, mineral lineation; Δ, computed maxima and minima of crystallographic directions.

massifs are known which present such a metamorphic sole¹⁸. The following description relies on structural studies conducted in six massifs and especially in the Bay of Island (Newfoundland)²⁰ and Samail (Oman)²¹.

Above the contact with the amphibolites and within a few hundred metres of it, the peridotites of the ophiolite suite

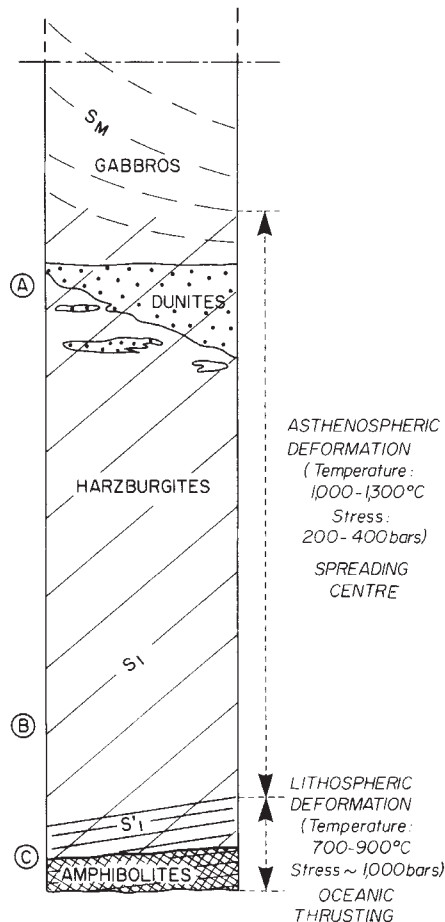


Fig. 3 a, Schematic cross-section through the lower part of an ophiolite sequence showing structures of the lithospheric deformation (S'_1) superimposed on that of the asthenospheric deformation (S_1). S_M is the cumulate layering. A, hypersolidus asthenospheric deformation; B, subsolidus asthenospheric deformation; C, lithospheric deformation. These zones refer to the models in Fig. 5.

exhibit mylonitic textures indicative of large strain-high stress deformation (Fig. 3). According to the olivine neoblast size piezometry, the stress is evaluated at 1–2 kbar (ref. 18). The temperature in these peridotites is estimated at 700–900 °C if they are considered to be in thermal equilibrium with the underlying amphibolites and granulites. This deformation is clearly distinct from the high temperature-low stress one (Fig. 3) which affected the whole ultramafic sequence and originated from the asthenospheric flow in a spreading environment. Temperatures during the asthenospheric flow are close to the peridotite solidus (1,250 °C) and stress is of the order of 200–300 bar (ref. 18).

The amphibolites underlying the peridotites often contain intensely deformed quartzite bands deriving from the silica-rich oceanic sediments. These rocks may have recorded several superimposed tectonic events corresponding to decreasing temperatures²². The highest temperature deformation, contemporaneous with the contact metamorphism, is ascribed to the oceanic thrusting. It produces a penetrative foliation and a lineation which are parallel with those in the overlying peridotites. The kinematic analysis of the thrusting can equally well be undertaken in the mylonitic peridotites or in the quartzite bands from the amphibolites. However, the low temperature limit for plastic deformation of quartz may mean that quartzites also recorded later events such as nappe emplacement.

Kinematics of oceanic thrusting

The proposed method has already been applied in the study of the Bay of Island ophiolite in Newfoundland²⁰ and of the Samail ophiolite in Oman²¹.

Figure 4 illustrates the case of the case of the Samail ophiolite. The high stress deformation traced at the very base of the peridotite has the same geometry as the deformation recorded in the amphibolites; it contrasts with the low stress-high temperature deformation producing coarse porphyroclastic textures in the main part of the peridotite sequence. Lattice preferred orientations of orthopyroxene (Fig. 4a) in mylonitic peridotite and of quartz (Fig. 4b) in quartzites display an obliquity to the penetrative foliation and lineation which indicate the same sense of shear for both rock types.

In geodynamic situations, two possible decoupling surfaces are envisaged (Fig. 5), in which thrusting of young oceanic crust could be initiated, due to oceanic convergence.

First, at a spreading axis itself²³, if there is a shift from oceanic expansion to convergence. The compressive force may initiate thrusting along the lithosphere–asthenosphere interface (Fig. 5a). The decoupling surface would correspond to the isotherm representing the lower limit for peridotite high temperature

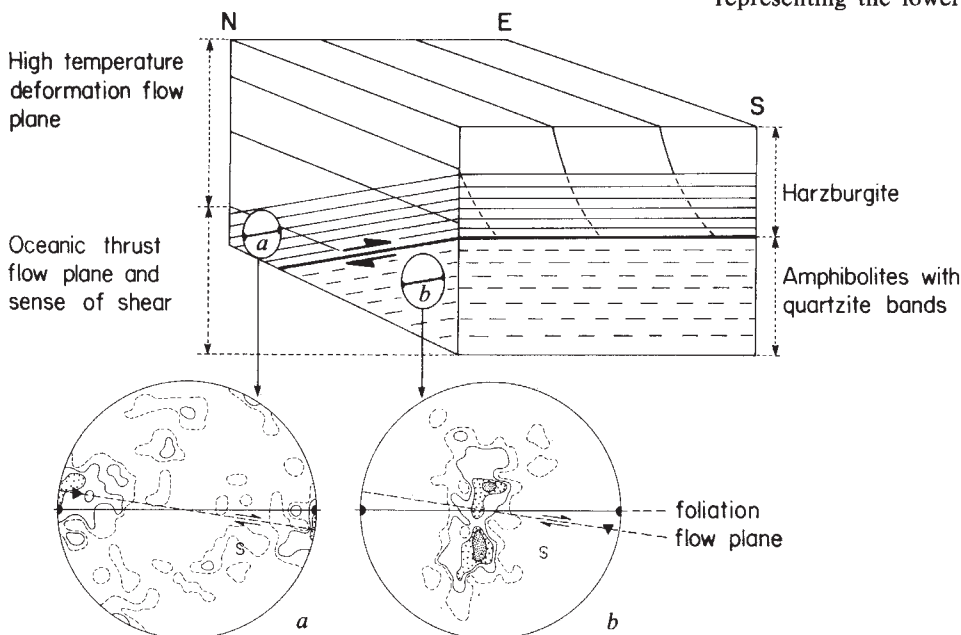


Fig. 4 Relationship between deformation in the amphibolites of the metamorphic sole and in the basal harzburgites in the Samail ophiolite (Oman). *a* and *b* are equal area lower hemisphere projections of crystallographic axes; 100 measurements; xz finite strain reference system (S = geographical south); solid line parallel to the lineation and trace of the foliation; the flow plane is deduced from fabric data; double arrow = sense of shear. *a*, [001] orthopyroxene slip direction in peridotite mylonite: $\blacktriangle$, the computed maximum (flow direction). *b*, c axis [0001] quartz in quartzite: $\blacktriangle$, the computed pole of the girdle (flow direction).

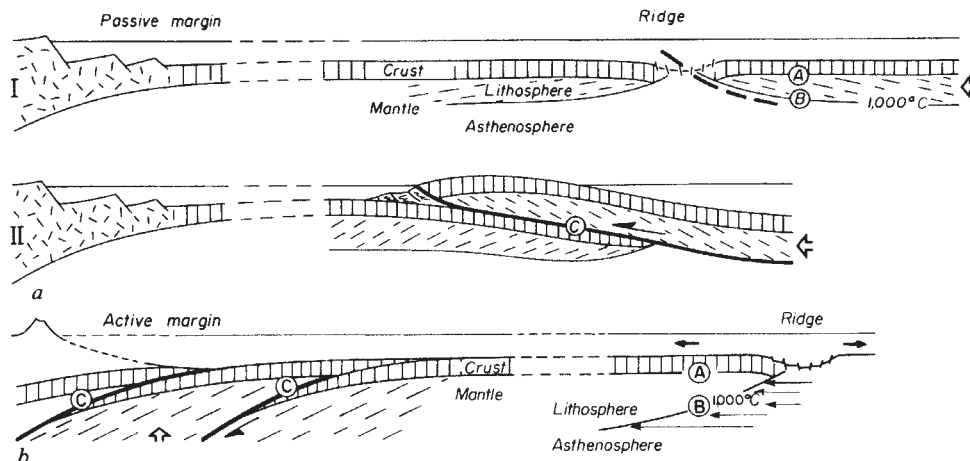


Fig. 5 Possible geodynamic locations of oceanic thrusting: *a*, at a spreading axis, along the 1,000 °C isograd, lithosphere/asthenosphere interface. A flat isograd corresponding to fast spreading will favour the expected situation: *b*, in a subduction environment at an active margin, at the time when new created crust is involved in the subduction process. A, B, C indicate zones where asthenospheric and lithospheric deformations are imprinted in the peridotite; it refers to level A, B, C, in the ophiolite sequence in Fig. 3.

plastic flow (>1,000 °C). Depending on the spreading rate, the isotherms are variously inclined (10°–25° for spreading rates 5–1 cm yr⁻¹). Such angles compare favourably with those proposed for basement thrust²⁴.

Second, at an active margin, where downward bending of the lithosphere plate develops elastic deviatoric stresses, seaward of the trench⁸. In the case of a young lithosphere, the

elastic stresses are located at shallow depth along a roughly horizontal decoupling surface.

In the first case, the thrusting will predate the establishment of a subduction process, as it is related with the initiation of the convergence process. In the second case the thrusting may postdate an active subduction, occurring at the time when young crust is involved in the subduction process.

Received 19 January; accepted 25 February 1982.

- Chase, C. G. *J. geophys. Res.* **83**, 5385 (1978).
- Le Pichon, X. *et al. Earth planet. Sci. Lett.* **44**, 441 (1979).
- Coleman, R. G. *J. geophys. Res.* **76**, 1212 (1971).
- Davies, H. *Colloq. Int. CNRS* **272**, 17 (1980).
- Aubouin, J., Mattauer, M. & Allegre, C. *C.r. hebdo. Séanc. Acad. Sci., Paris* **285**, 953 (1977).
- Malpas, J., Stevens, R. & Strong, D. F. *Geology* **1**, 45 (1973).
- Parrot, J. F. & Whitechurch, H. *Rev. Geogr. Phys. Geol. Dyn.* **20**, 153 (1978).
- Nicolas, A. & Le Pichon, X. *Earth planet. Sci. Lett.* **46**, 597 (1980).
- Nicolas, A., Bouchez, J. L., Boudier, F. & Mercier, J. C. *Tectonophysics* **12**, 55 (1971).
- Nicolas, A. & Poirier, J. P. *Crystalline Plasticity and Solid State Flow in Metamorphic Rocks* (Wiley, New York, 1976).
- Bouchez, J. L. *Tectonophysics* **49**, 25 (1978).

- Bouchez, J. L. *Tectonophysics* **39**, 25 (1977).
- Etchecopar, A. *Tectonophysics* **39**, 121 (1977).
- Bouchez, J. L. & Pecher, A. *Tectonophysics* **78**, 23 (1981).
- Hudleston, P. J. *Energetic of Geological Process*, 121 (Springer, Berlin, 1977).
- Burg, J. P. & Laurent, P. *Tectonophysics* **47**, 15 (1978).
- Bouchez, J. L. & Duval, P. *Terra Cognita Spec. Iss.*, 60 (1981).
- Nicolas, A., Boudier, F. & Bouchez, J. L. *Am. J. Sci.* **20**, 153 (1978).
- Williams, H. & Smyth, W. R. *Am. J. Sci.* **273**, 594 (1973).
- Girardeau, J. & Nicolas, A. *Tectonophysics* **77**, 1 (1981).
- Boudier, F. & Coleman, R. G. *J. geophys. Res.* **86**, 2573 (1981).
- Jamieson, R. A. *J. Petrol.* **22**, 397 (1981).
- Boudier, F. & Michard, A. *Terra Cognita* **1**, 109 (1981).
- Elliot, D. *Phil. Trans. R. Soc. A* **283**, 289 (1976).

Enzymatic synthesis of bacteriophage fd viral DNA

Thomas F. Meyer & Klaus Geider

Max-Planck-Institut für medizinische Forschung, Abteilung Molekulare Biologie, Jahnstrasse 29, D 6900 Heidelberg, FRG

An enzyme system with requirements similar to those for replication of phage fd replicative form (RF) DNA in bacteriophage fd-infected cells has been reconstituted with purified fd gene 2 protein, and DNA polymerase III holoenzyme, DNA binding protein I and rep-protein (rep-helicase) of Escherichia coli. The system generates viral circular single strands, which are infective for E. coli spheroplasts. Parental and newly synthesized DNA are covalently connected in early stages of replication, as expected for DNA replication using the rolling circle mechanism. Single-stranded tails of the rolling circle intermediates are cleaved after a full round of replication by gene 2 protein and circularized by the same enzyme molecule.

THE F-pili specific filamentous bacteriophage fd, which is closely related to phages M13 and f1, has been extremely useful in the study of structural and functional aspects of small viral DNA genomes. The organization of the circular fd genome and its nucleotide sequence have been analysed in detail, and the single-stranded character of fd DNA has given rise to the construction of unique vector systems for cloning and quick sequencing of genes¹.

The life cycle of bacteriophage fd² begins with the entry of the viral DNA into an *Escherichia coli* cell, where it is immediately converted into a double-stranded replicative form. RF replication then occurs, forming a pool of double-stranded DNA molecules. RF replication requires the phage-encoded gene 2 protein as the sole enzyme function encoded by the phage³. In a late stage of infection the pool of RF molecules

is used as a matrix for the production of new viral single strands.

On the basis of genetic data, host functions of *Escherichia coli* essential for replication of filamentous phages have been assigned to DNA polymerase III and I⁴, RNA polymerase, DNA gyrase and probably DNA binding protein I⁵. The involvement of *E. coli* RNA polymerase and DNA gyrase has been deduced from the inhibition of phage replication by the drugs rifampicin⁶ and nalidixic acid⁷. Rifampicin specifically interferes with the synthesis of the complementary strand in the stage of duplex propagation⁸, whereas nalidixic acid affects RF replication, but not the conversion of single strands to RF⁷. Another *E. coli* gene function involved in RF replication of small bacteriophages is encoded by the *rep* gene⁹. Participation of the *dnaB* and *dnaG* gene products has also been proposed

for replication of filamentous phages^{10,11}, although their requirement in enzymatic replication of fd RF has not been demonstrated.

Enzymatic studies have revealed the participation of the *E. coli* proteins DNA binding protein I, RNA polymerase, DNA polymerase III holoenzyme, DNA polymerase I and DNA ligase in the conversion of single strands to RF¹². The polymerases are required, respectively, for synthesis of the RNA primer, elongation of the DNA chain and excision of the primer concomitant with gap filling¹³. The DNA binding protein directs the initiation event and supports DNA polymerase III in DNA replication. DNA ligase seals the double-stranded DNA into the relaxed circular covalently closed form (RFIV).

To initiate double-strand replication, the relaxed RF is converted into supercoiled RFI by DNA gyrase. Phage fd gene 2

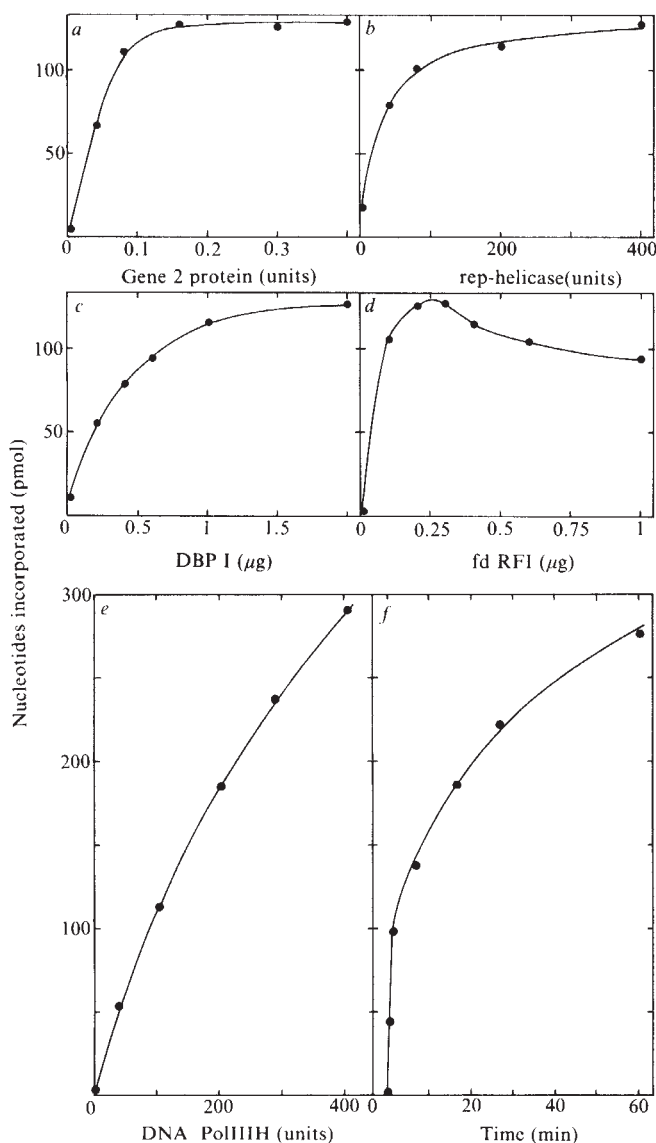


Fig. 1 Proteins required for enzymatic replication of fd RF. The standard incubation mixture contained 200 ng fd RFI (600 pmol of nucleotides), 0.08 U of fd gene 2 protein, 100 U of *E. coli* rep-helicase, 1 μ g *E. coli* DNA binding protein I, 100 U of *E. coli* DNA polymerase III holoenzyme in 20 μ l of 20% sorbitol, 50 mM HEPES (pH 7.5), 0.1 mg ml⁻¹ bovine serum albumin, 5 mM 2-mercaptoethanol, 5 mM MgCl₂, 500 μ M ATP, 50 μ M of dATP, dGTP and dTTP, and 25 μ M of ³H- or ³²P-labelled dCTP (1,000–3,000 c.p.m. pmol⁻¹). Incubation was for 5 min at 30 °C. The nucleotide incorporation was stopped with 5% perchloric acid and the acid-insoluble radioactivity determined. The sources of the proteins have been described elsewhere¹⁷. *a*, fd gene 2 protein; *b*, *E. coli* rep-helicase; *c*, *E. coli* DNA binding protein I; *d*, phage fd RFI; *e*, *E. coli* DNA polymerase III holoenzyme. *f* Gives the time course of nucleotide incorporation.

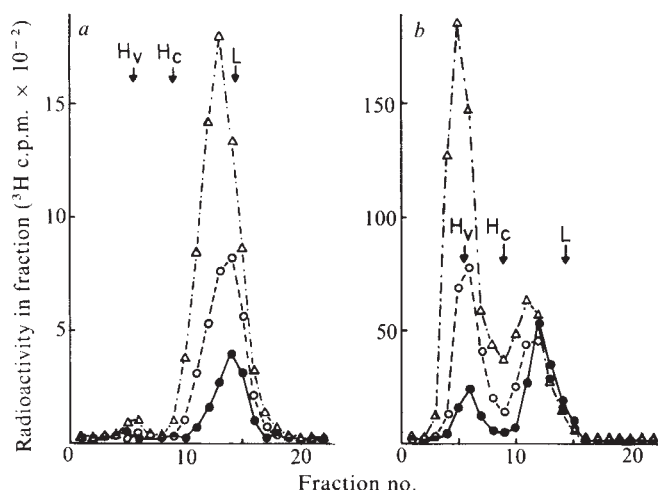


Fig. 2 Analysis of bromouracil-labelled replication products in alkaline caesium sulphate. In the standard incubation mixture, dTTP was replaced by bromodeoxyuridine triphosphate. Incubation was stopped at various times by the addition of 80 μ l of 0.1 M NaOH and 10 mM EDTA. The Cs₂SO₄ gradients were run in a Beckman 50 Ti fixed-angle rotor at 40,000 r.p.m. and at a mean density of 1.46 g ml⁻¹ for 28 h at 20 °C in 0.08 M NaOH/10 mM EDTA. Increasing density is from right to left. Panel *a*: ●, 15 s; ○, 30 s; △, 60 s. Panel *b*: ●, 2 min; ○, 4 min; △, 8 min of incorporation. External markers were bromodeoxyuridine-substituted fd viral strands (H_v) and complementary strands (H_c) and fd single-stranded DNA of light density (L).

protein cleaves this DNA in the viral strand¹⁴ at a specific site^{15,16}. Gene 2 protein supports strand unwinding by rep-helicase (rep-protein)¹⁷ and the 3'-hydroxy end of the viral strand is extended by DNA polymerase III holoenzyme. Both reactions require the presence of *E. coli* DNA binding protein I. Cleavage and circularization of replicated viral strands by fd gene 2 protein have been demonstrated in model systems^{18,19}. In addition, replication of duplex DNA from filamentous phages has been reported for a subcellular *E. coli* system²⁰.

Information on the mechanism of DNA replication has been obtained with Φ X174 DNA for the conversion of viral strands into supercoiled duplex DNA²¹, and its simultaneous replication into viral and complementary strands²². Although this system is structurally related to the phage fd replication system, it differs substantially regarding enzymatic requirements and the mode of action of the enzymes involved^{13,23}. Replication systems using linear double-stranded DNA as a substrate, that is, the T4²⁴ and the T7²⁵ *in vitro* systems, have given insight into the mechanistic principles of a replication fork.

Here, we analyse an enzyme system for replication of double-stranded fd DNA which is assembled with proteins present in the fd-infected cell. We will discuss in detail the mode of action of gene 2 protein as the key enzyme in the production of circular bacteriophage fd viral strands.

Enzymatic requirements for replication of fd RFI

In accordance with the requirements for phage fd replication *in vivo*, an enzymatic system was set up for components stimulating DNA synthesis on supercoiled fd RFI DNA. The influence of the individual proteins—fd gene 2 protein, *E. coli* rep-helicase, *E. coli* DNA binding protein I and *E. coli* DNA polymerase III holoenzyme—and of the template DNA fd RFI on the incorporation of deoxyribonucleotides is shown in Fig. 1. The DNA and the replication proteins can be added up to saturating amounts in standard incubation conditions, except that increasing amounts of DNA polymerase III holoenzyme do not reach a plateau of nucleotide incorporation (Fig. 1e). DNA synthesis starts immediately after incubating the assay mixture at 30 °C, decreases slightly after 5 min of incubation,

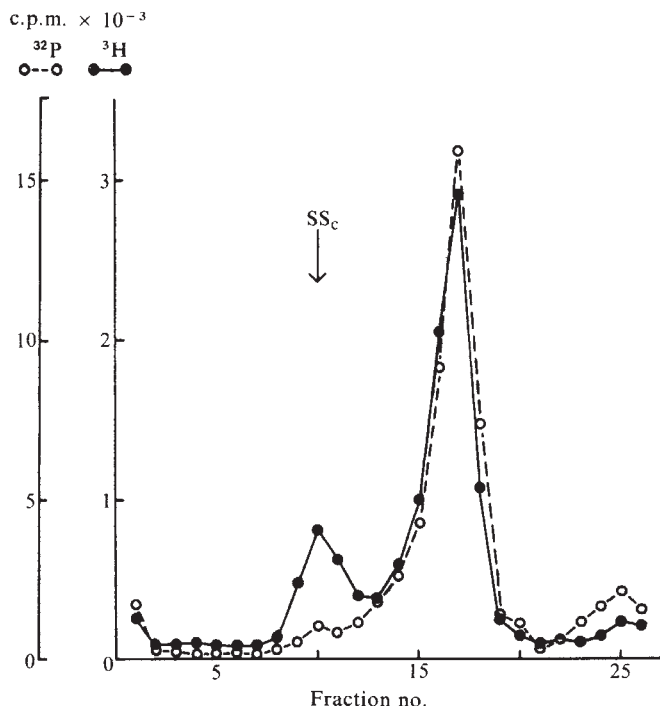


Fig. 3 Gradient analysis of replication products. *a*, Sedimentation in neutral sucrose. The template ^3H -fd RFI (10,000 c.p.m., 300 ng) was isolated from phage-infected cells. A volume of 60 μl (three times the assay system described in Fig. 1 legend) with label in $[\alpha\text{-}^{32}\text{P}]\text{dCTP}$ (1,700 c.p.m. pmol^{-1}) was incubated for 5 min at 30 $^{\circ}\text{C}$. After the reaction had been stopped with NaCl/EDTA, the sample was diluted with 100 μl of 1 M NaCl in Tris-HCl/EDTA and layered on top of a neutral sucrose step gradient. Sedimentation was performed in a Beckman SW60 Ti rotor for 2.5 h at 50,000 r.p.m. and 20 $^{\circ}\text{C}$. To fractionate the gradient the tubes were punctured in the bottom. Sedimentation is from right to left. The position of fd circular single strands (SS_C) as external marker is indicated.

and is still occurring at 60 min (Fig. 1f). At that time, the amount of nucleotides incorporated approaches the equivalent of one round of viral strand synthesized per RF molecule.

Products generated in the replication system

To determine the mode of replication implemented in the enzyme system, density label was incorporated into newly synthesized DNA via substitution of thymidine triphosphate by bromodeoxyuridine triphosphate. Subsequent equilibrium sedimentation analysis of the replicated DNA in denaturing conditions reveals label in two distinct positions: at heavy density, corresponding to heavy viral marker DNA, and at intermediate-light density (Fig. 2). The latter material, which is predominantly found at early incubation times, indicates a covalent connection of heavy newly synthesized and light parental DNA. These results support the rolling circle model for fd double-strand replication which implies synthesis of the viral strand by covalent elongation of the parental viral DNA. The tail of single-stranded DNA generated in this reaction is then cleaved off late in the replicative cycle. Heavy material in the density gradient represents this kind of processed viral DNA which is found after ~60 s of incubation.

To monitor the production of fd single strands, ^3H -labelled fd RFI was used as template and newly synthesized material was labelled with ^{32}P -dCTP. Sedimentation of the products in neutral sucrose separated single-stranded from double-stranded DNA (Fig. 3). Quantitative evaluation of the distribution of both labels revealed that 25% of the total radioactivity of the template occurred as single-stranded DNA. About 6% of the newly synthesized ^{32}P label sedimented in the position of single-stranded material. These results show that half of the template molecules were replicated and the single strands with

parental label were cleaved from the double-stranded core, but label from the second round of replication could not be detected efficiently by sedimentation in neutral sucrose.

We have further investigated the replication products by electrophoresis in agarose gels, where linear and circular single strands can be separated from each other. Figure 4a shows that the single-stranded newly synthesized DNA migrated in the position of circles, whereas linear single strands of unit length were less than 1% of the circular DNA. This result illustrates the efficient circularization of single-stranded DNA by gene 2 protein in the *E. coli* replication system. Because no label is incorporated into single strands in the first round of replication, these circles are derived from following rounds of viral strand synthesis.

The production of circular fd single strands was shown by migration with regard to single-stranded marker DNA (Fig. 4a), but also proven by conversion of the material into double-stranded DNA (Fig. 4b, c). The enzymatically synthesized single-stranded DNA was nonspecifically primed with plasmid-encoded DNA primase²⁶, the RNA primer was extended with

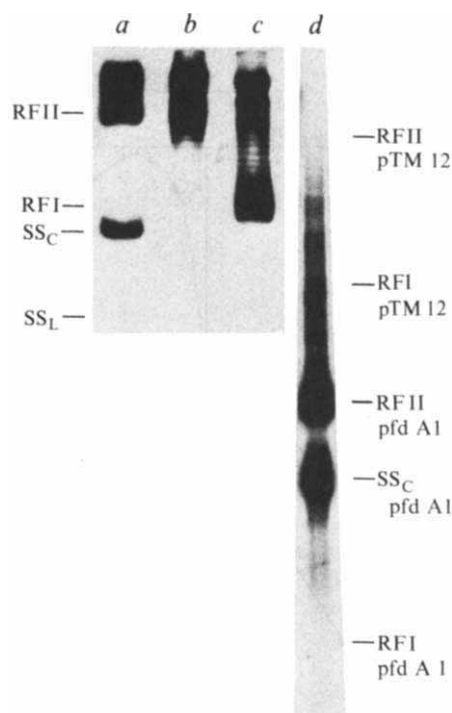


Fig. 4 Electrophoresis of replication products on agarose. *a*, A standard incubation mixture (see Fig. 1) with $[\alpha\text{-}^{32}\text{P}]\text{dCTP}$ (720 c.p.m. pmol^{-1}) was stopped with SDS buffer and the DNA separated on a 1% agarose gel. *b*, Incubation was as in *a* but without label. The mixture was then heated to 65 $^{\circ}\text{C}$ for 5 min and incubated after addition of $[\alpha\text{-}^{32}\text{P}]\text{dCTP}$ with DNA primase encoded by the I-like plasmid R64drd11 (from *E. Lanka*)²⁶, *E. coli* DNA polymerase I (0.1 μg) and T4 DNA ligase (2 μg)³ for 1 h at 20 $^{\circ}\text{C}$. *c*, The products of *b* were incubated with 2 units of purified DNA gyrase¹⁷ for 30 min at 30 $^{\circ}\text{C}$. External markers were single-stranded circular fd DNA (SS_C) isolated from the purified phage, linear single-stranded fd DNA (SS_L) obtained by cleavage of fd single-stranded circles with pancreatic DNase; fd RFI (RFI) isolated from fd-infected cells; fd RFII_o (RFII) generated by cleavage of RFI with fd gene 2 protein and linear single strands (RFIII), obtained by cleavage of fd RFI with gene 2 protein in the presence of Mn^{2+} (ref. 14). *d*, Replication of plasmid pfdA1 (2.4 kb)³. The plasmid (200 ng), which contains a cloned 261-base pair origin fragment of phage fd, was replicated in conditions described in Fig. 1 legend. In addition, 50 ng of the pBR derivative pTM12 (6.2 kb)³, bearing the gene 2 of fd but not the fd origin, were present in the assay and served as an internal control for the specificity of the replication system. The products were separated by electrophoresis through 1.5% agarose. The single-stranded nature of the pfdA1 replication product, SS_C , was controlled by digestion with S_1 nuclease.

Table 1 Gene 2 protein-mediated activation of fd RF replication

Substrate	DNA synthesis by the four-component system (pmol)		Product of first round of replication
	-G2P	+G2P	
fd RFI	5	120	SS circles
fd RFIV	6	3	—
fd RFII ₀	18	84	SS circles
fd RFII _p	14	103	SS linear DNA of unit length
fd RFII	16	15	—

The template DNA (200 µg) was added to the standard assay system (see Fig. 1). Products were analysed by agarose gel electrophoresis and subsequent quantification of unlabelled (first round of replication) and labelled (following rounds of replication) single-stranded (SS) DNA species. The analysis was accompanied by electron microscopic evaluation¹⁷. G2P, gene 2 protein; fd RFII₀, phage fd replicative form DNA specifically cleaved with fd gene 2 protein; RFII_p represents RFII, without 5'-phosphate on the cleaved viral strand.

DNA polymerase I and the strands were finally sealed with T4 DNA ligase (Fig. 4b). Subsequent treatment of this DNA with gyrase should shift all closed circular molecules into the position of supercoils in the gel. This indeed can be observed in Fig. 4c, proving the circular nature of the originally synthesized single strands. Control experiments in the absence of primase showed that an excess of T4 DNA ligase prevented incorporation of label into double-stranded fd RFII, which was also produced in the replication system.

Enzymatic replication with DNA species different from fd RFI

The initiating cleavage reaction conducted by gene 2 protein is limited on supercoiled DNA substrates which harbour the replication origin of phage fd¹⁴. The small 2.4-kilobase (kb) plasmid pfdA1, which contains a cloned fragment of the fd origin³, should fulfill these properties. Figure 4d shows that the described *in vitro* system uses the 261-base pair origin fragment and replicates the plasmid with high efficiency, producing single-stranded DNA.

Previous studies on the mechanism of gene 2 protein revealed a specific reversible binding affinity of the enzyme to its cleavage site in fd RFII₀. This weak interaction is an essential prerequisite for unwinding of fd RFII₀ by *rep*-helicase and DNA binding protein I¹⁷. Consequently, fd RFII₀ is a suitable substrate in the replication system for production of circular single strands like fd RFI, if gene 2 protein is present (Table 1). Removal of the 5'-phosphate group in fd RFII₀ still produces a substrate (fd RFII_p) which is efficiently propagated by gene 2 protein and the *E. coli* proteins. However, linear rather than circular single strands are generated in this case due to a missing phosphate group required for ring closure in the first round of replication. Relaxed RFIV, and fd RFII with a nick outside the replication origin, were unsuitable templates in the replication system composed of the four proteins (Table 1).

Biological activity of the replicated single-stranded circles

Faithful DNA replication should produce DNA with the informational content of the template DNA; imprecise circularization of viral single strands should result in a loss of their biological activity. We analysed the *in vitro* replication products in an *E. coli* protoplast system in which double-stranded phage DNA is ~100 times less infective than single-stranded circular DNA (Table 2).

The replication products from a standard incubation assay were deproteinized, because *E. coli* DNA binding protein I interferes with the infectivity of single strands²⁷. Plaque formation in the spheroplast assay (Table 2) showed infective centres

after an enzymatic replication of the DNA, which exceeded infectivity of double-stranded fd RF by two orders of magnitude. Single strands produced by strand unwinding of RF increased this background, but linear viral strands or circular complementary strands were apparently less infective than circular viral strands, as already found for phage ΦX174 DNA²⁸. These results demonstrate effective fd viral strand replication in our system and the precise circularization of the synthesized viral strands by gene 2 protein.

The role of fd gene 2 protein in phage duplex DNA replication

Replication of the viral strand of fd RF is initiated by the phage-encoded gene 2 protein^{29,30}, which introduces a specific single-strand break into the viral strand of the phage RF. The cleavage site of gene 2 protein has been localized in a region of palindromic symmetry¹⁵. This finding and the observation that gene 2 protein cuts both strands of fd RF in the presence of Mn²⁺ (ref. 14) suggest that in physiological conditions two molecules of the enzyme bind to the cleavage region, but only one reacts with the viral strand. Consequently, dimer formation at the cleavage site is believed to be essential in the initial step, and also for processing of rolling circle structures (see below). Concomitant with the cleavage of the viral strand, gene 2 protein may attach transiently to one of the termini of this strand by forming a covalent bond. This intermediate may provide the energy to seal the previously opened strand in a topoisomerase-like manner¹⁴. The attachment of the enzyme to the nicked DNA strand is either followed by the formation of RFIV or the hypothesized protein-DNA complex is

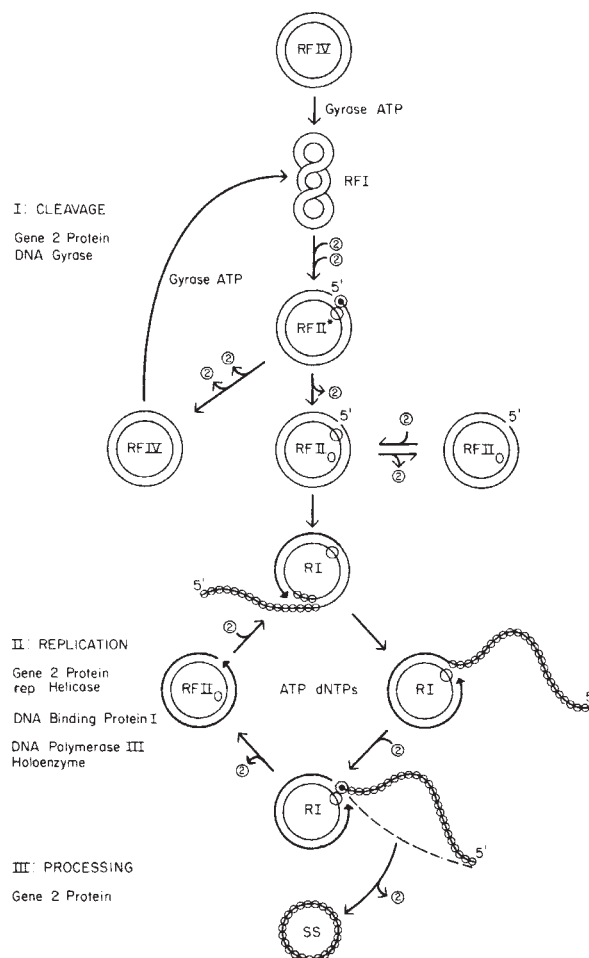


Fig. 5 A model for enzymatic replication of phage fd RF and the involvement of bacteriophage fd gene 2 protein.

Table 2 Infectivity of newly formed fd single strands after enzymatic replication of phage fd RF

DNA source	Infective centres per 1×10^{10} molecules of template DNA
Replication of fd RF	2,780
Enzymatic unwinding of fd RF*	400
Cleavage of fd RFI by fd gene 2 protein†	5
fd RFI‡	47
fd viral single strands	8,500

Phage fd RF was replicated for 20 min as described in Fig. 1 legend. The incubation was stopped after 5 min at 30 °C by the addition of EDTA. 0.05 M Tris-HCl (pH 8.0) was added to a final volume of 0.3 ml and the sample was treated with phenol. Phenol in the aqueous phase was removed by extraction with ether. The sample was then mixed with 0.2 ml of a protoplast suspension of cells from *E. coli* K-12 strain W6 (ref. 33). After 15 min at 37 °C the protoplasts were mixed with the indicator strain 1101 and plated on PAM agar.

* Unwinding with fd gene 2 protein, *rep*-helicase and DNA binding protein I. The products are circular complementary strands and linear viral strands.

† The reaction with gene 2 protein gives RFIV and RFII₀ (ref. 14).

‡ Transfection rate variable.

rearranged to become active for initiation of the unwinding of fd RFII₀.

Analysis of the cleavage site in fd RFII has revealed free 3'-hydroxy and 5'-phosphate ends which are substrates for ligation with DNA ligase or for nick-translation with DNA polymerase I^{14,15,31}. However, the complementary strand at the cleavage site seems to be preferentially protected against endonucleolytic digestion with *Bal*31 nuclease, in contrast to randomly located nicks¹⁷. It has been shown that this protection is due to a reversible binding of gene 2 protein to the cleavage site, which is necessary for specific initiation of the unwinding and the replication reactions. RFIV which is produced in the initiation reaction in addition to fd RFII₀ needs to be recycled into RFI by DNA gyrase before being processed once more by gene 2 protein (Fig. 5).

The subsequent replication process, outlined in Fig. 5, requires the *E. coli* proteins DNA polymerase III holoenzyme, *rep*-helicase and DNA binding protein in addition to the fd gene 2 protein. The *rep*-helicase accomplishes unwinding of fd RFII in collaboration with DNA binding protein I¹⁷. DNA polymerase III holoenzyme extends the 3'-hydroxy end released by the cleavage reaction and proceeds in a rolling circle-type replication. As deduced from experiments carried out in the phage T4 *in vitro* replication system¹⁹, one molecule of gene 2 protein remains bound to the complementary strand of fd RF during replication. As replication ceases near the origin in the presence of gene 2 protein, this enzyme may serve as a signal for termination after a complete round of replication. A second molecule of gene 2 protein is then required to cut off the replicated single strand, based on the difference in the kinetics of both steps¹⁹. The parental RF molecule is released in an open form and therefore immediately used as a template

for further replication cycles. For late infection with filamentous phages the occurrence of pulse label in the viral strand of RFIV³² may be due to conversion of RFII molecules into doubly closed DNA by DNA polymerase I and DNA ligase.

The unwound single strands are readily covered by DNA binding protein as replication proceeds¹⁷. This stoichiometric function can be substituted by phage-encoded DNA binding protein (fd gene 5 protein), although catalytic amounts of DNA binding protein I remain essential for unwinding and synthesis of DNA (unpublished results).

In a final reaction, the unwound viral single strand is separated from the parental RF molecule and circularized (Fig. 5). The function of gene 2 protein is solely responsible for this process, as the *E. coli* proteins in the replication system can be substituted completely by T7 enzymes, that is, T7 DNA polymerase and T7 DNA primase (gene 4 protein)¹⁸. However, as we demonstrated by use of the T4 replication system, processing only occurs during concomitant DNA synthesis¹⁹. In the absence of enzymes providing DNA synthesis, long-tailed rolling circle structures are not processed *per se*, which is consistent with the observation that circular viral single strands are not accepted as a substrate by gene 2 protein¹⁴. We therefore believe that, as for cleavage of fd RFI, a defined DNA conformation in the cleavage region is required for processing the rolling circle intermediate. This situation may be realized by a partially melted double-stranded structure caused by the superhelical tension in fd RFI or due to unwinding, when the replication fork enters the cleavage region of the rolling circle intermediate¹⁹.

The circularization of phage fd DNA by gene 2 protein basically resembles the sequence-specific topoisomerase reaction of this enzyme on fd RFI. Although we cannot demonstrate the independence on ATP of this reaction, because strand unwinding requires ATP, we have shown previously that the nicking-closing reaction of gene 2 protein on fd RF is unaffected by high-energy compounds like ATP¹⁴. As cleavage of single-stranded tails from the rolling circle intermediate occurs regardless of the composition of the 5' end of the tail¹⁹ (Table 1), we conclude that gene 2 protein attaches to the 3'-hydroxy end so as to conserve the cleavage energy for circularization of this DNA. Some affinity of the 5' end at the tail of the rolling circle for the gene 2 protein molecule on the complementary strand may favour the replicated viral strand looping back to the double-strand core, thus facilitating viral strand circularization after a round of replication. Circularization may also be supported by pairing palindromic sequences at the replication origin.

If the 5'-terminal phosphate group of the cleaved viral strand is missing, no circle can be formed (Table 1)¹⁷ and gene 2 protein dissociates from the DNA. Thus, gene 2 protein can act in two different ways: either it uses the conserved chemical energy to form a new phosphodiester bond, or, as for the production of RFII₀, it leads to a decay of the protein-DNA complex, which then simply mimics an endonucleolytic reaction.

We thank Erich Lanka for the donation of DNA primase, Marlies Kaiser, Erna Piaskowski and Ilge Wilhelm for technical assistance and Hermann Bujard, David H. Figurski and Robb Moses for comments on the manuscript.

Received 10 December 1981; accepted 1 March 1982.

- Beck, E. & Zink, B. *Gene* **16**, 35–58 (1981).
- Hohn, B. & Marvin, D. *Bact. Rev.* **33**, 172–209 (1969).
- Meyer, T. F. & Geider, K. *Proc. natn. Acad. Sci. USA* **78**, 5416–5420 (1981).
- Staudenbauer, W. *Eur. J. Biochem.* **49**, 249–256 (1974).
- Glassberg, J., Meyer, R. R. & Kornberg, A. *J. Bact.* **140**, 14–19 (1980).
- Brutlag, D., Schekman, R. & Kornberg, A. *Proc. natn. Acad. Sci. U.S.A.* **68**, 2826–2829 (1971).
- Schneck, P. K., Staudenbauer, W. L. & Hofschneider, P. H. *Eur. J. Biochem.* **38**, 130–136 (1973).
- Fidanián, H. M. & Ray, D. S. *J. molec. Biol.* **83**, 63–82 (1974).
- Denhardt, D. T. *Comprehensive Virol.* **7**, 1–104 (1977).
- Olsen, W. L., Staudenbauer, W. L. & Hofschneider, P. H. *Proc. natn. Acad. Sci. U.S.A.* **69**, 2570–2573 (1972).
- Ray, D. S., Dueber, J. & Suggs, S. *J. Virol.* **16**, 348–355 (1975).
- Geider, K. & Kornberg, A. *J. biol. Chem.* **249**, 3999–4005 (1974).
- Meyer, T. F. & Geider, K. *ICN-UCLA Symp. molec. cell. Biol.*, 579–588 (1980).
- Meyer, T. F. & Geider, K. *J. biol. Chem.* **254**, 12642–12646 (1979).
- Meyer, T. F., Geider, K., Kurz, C. & Schaller, H. *Nature* **278**, 365–367 (1979).
- Horiuchi, K. & Zinder, N. D. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2341–2345 (1976).

- Geider, K., Bäuml, I. & Meyer, T. F. *J. biol. Chem.* **257** (in the press).
- Harth, G., Bäuml, I., Meyer, T. F. & Geider, K. *Eur. J. Biochem.* **119**, 663–668 (1981).
- Meyer, T. F., Bäuml, I., Geider, K. & Bedinger, P. *J. biol. Chem.* **256**, 5810–5813 (1981).
- Schneck, P. K., van Dorp, B., Staudenbauer, W. L. & Hofschneider, P. H. *Nucleic Acids Res.* **5**, 1689–1700 (1978).
- Shlomai, J., Polder, L., Arai, K.-I. & Kornberg, A. *J. biol. Chem.* **256**, 5233–5238 (1981).
- Arai, N., Polder, L., Arai, K.-I. & Kornberg, A. *J. biol. Chem.* **256**, 5239–5246 (1981).
- Shlomai, J. *et al. ICN-UCLA Symp. molec. cell. Biol.*, 545–568 (1980).
- Alberts, B. M. *et al. ICN-UCLA Symp. molec. cell. Biol.*, 449–473 (1980).
- Tamanoi, F. *et al. ICN-UCLA Symp. molec. cell. Biol.*, 411–428 (1980).
- Lanka, E., Scherzinger, E., Günther, E. & Schuster, H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3632–3636 (1979).
- Uhlmann, A. & Geider, K. *Biochim. biophys. Acta* **474**, 639–645 (1977).
- Goulian, M., Kornberg, A. & Sinsheimer, R. L. *Proc. natn. Acad. Sci. U.S.A.* **58**, 2321–2328 (1967).
- Lin, N. S.-C. & Pratt, D. *J. molec. Biol.* **72**, 37–49 (1972).
- Fidanián, H. M. & Ray, D. S. *J. molec. Biol.* **72**, 51–63 (1972).
- Meyer, T. F. & Geider, K. *J. biol. Chem.* **254**, 12636–12641 (1979).
- Horiuchi, K., Ravetch, J. V. & Zinder, N. D. *Cold Spring Harbor Symp. quant. Biol.* **43**, 389–399 (1978).
- Guthrie, G. D. & Sinsheimer, R. L. *Biochim. biophys. Acta* **72**, 290–297 (1963).

LETTERS TO NATURE

Transient emission of ultra-high energy pulsed γ rays from Crab pulsar PSR0531

A. I. Gibson, A. B. Harrison, I. W. Kirkman,
A. P. Lotts, J. H. Macrae, K. J. Orford, K. E. Turver
& M. Walmsley

Physics Department, University of Durham, Durham DH1 3LE, UK

A new experiment has been established to measure the energy spectrum of ultra-high energy (UHE) γ rays ($E > 2,000$ GeV) from a number of celestial objects. During the commissioning of the equipment several potential sources were studied, including the Crab pulsar PSR0531. UHE γ rays from this object have been observed in previous experiments to be practically all pulsed¹⁻⁴, but no single observation has been of high statistical significance, and the flux values have been difficult to reproduce⁵. The current observations are based on 34 h of exposure between 25 September and 2 November 1981, and show overall an integral pulsed flux in agreement with our estimate of the average of previous data; most important, they also show strong evidence for a burst of pulsed emission of ~ 15 min duration. There had been previous suggestions that the Crab pulsar is variable at the energy we used and at slightly lower energies^{1,3}, but at time scales of months or days. The present work is the first to demonstrate short duration (15 min) emission and to offer an explanation of previous apparently discordant results.

The details of the experiment are given elsewhere⁶; it comprises an array of four independent telescopes designed to detect the weak optical Cherenkov emission from electron showers produced in the atmosphere by cosmic γ rays or cosmic-ray particles. It is sited at an altitude of 1,451 m in the Great Salt lake desert, USA. Each telescope contains three paraxial mirrors of 1.5 m diameter focusing the Cherenkov light onto photomultipliers with a field of view of 1.7° FWHM. The array is designed to enable sources to be continuously tracked, with the arrival directions of cosmic γ rays or particles entering the field of view being measured by using the relative time of arrival of the Cherenkov light front at each telescope. When searching for point sources of continuous emission, the detectors are operated with procedures designed to stabilize the threshold energy⁶, which results in an elevated threshold and consequently a lower shower detection rate.

The results reported here are based principally on the universal time (UT) of arrival of cosmic-ray events in the field of view. The search for periodic emission requires the highest possible rate of shower detection, and so the stability procedures were not employed, in order to reduce the threshold energy. The times of arrival of UHE γ rays and the much more numerous background cosmic-ray particles were recorded with a resolution of $1 \mu\text{s}$, and with an absolute uncertainty in UT of ~ 1 ms. After reduction to the Solar System's barycentre, the times were unfolded using an ephemeris derived from contemporary radio observations (A. G. Lyne, personal communication) (period = 0.033263833076 s, derivative = 36.416 ns day⁻¹, epoch = 2444910.5). The phase histograms so produced were searched for evidence of non-uniformity, arising from pulsed γ rays being detected at a uniform rate over the period of observation. Overall, no significant signal was found.

Ideally, periodicity should be searched for by correlating the observed γ -ray arrival time series with the expected light curve, using a likelihood ratio method. This cannot, however, be used in this case because there are no well measured data on the

light curve at these energies. There are only estimates based on measurements at lower energies, to which *a priori* probabilities cannot be assigned. The usual procedure then adopted is to invoke Poincaré's Theorem—that a random time series of sufficient length, unfolded with an arbitrary period, results in a uniform phase distribution. A test for the uniformity of the phase distribution, resulting from the experimental time series being unfolded with a trial period, is taken to indicate the absence of that period in the time series. Previous searches for periodicity have usually used a binned phase distribution and then either Pearson's χ^2 test for uniformity, or have searched by eye for evidence of an excess in one or more bins, with an allowance for the extra degrees of freedom (d.f.) arising from the arbitrary choices. These methods suffer from some disadvantages due to the binning and the arbitrary choices to be made. However, more importantly, none of these constitutes the most powerful test, even if independent tests are combined. The 'uniformly most powerful' test for the uniformity of a circular (phase) distribution, which also avoids the problem of arbitrary choice, is the Rayleigh test⁷ in which the resultant vector, Φ , formed from unit vectors at the phases of the events in the time series is tested for significance. The statistic $2N\Phi^2$ (where N is the number of events in the time series) is distributed as a χ^2 distribution with 2 d.f. This is equivalent to the test for the significance of an amplitude in power spectral analysis, in which the amplitude of any period from a noise series also has a χ^2 distribution with 2 d.f. (ref. 8). This procedure has been extensively checked by testing randomly generated time series, both with and without an added sampled periodic component. The distribution of the resultant, Φ , was found to conform closely to that predicted in the case of such noise series, and for randomly chosen periods in the experimental series. It has been possible to verify the distribution only down to a probability of 10^{-5} due to the limitations of the

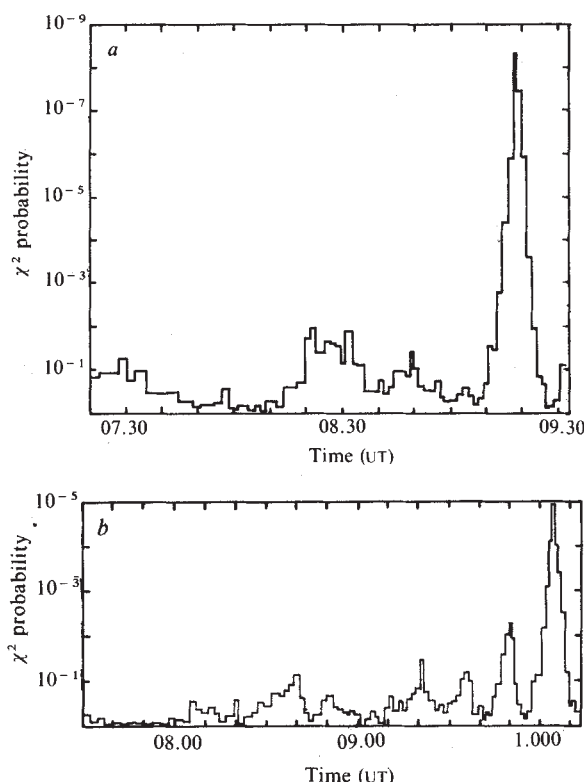


Fig. 1 The probability of lack of periodicity at the radio period for data in 15 min intervals during observations on *a*, 23 October 1981; *b*, 31 October 1981.

pseudo-random number generator used, and to the finite length of our experimental data. An added periodic component was always detected at the injected strength but with a sampling uncertainty given by the above distribution.

Random sampling of a noise series incorporating a true periodicity should give, and in our simulations gave, an effect close to the nearest Fourier harmonic of the duration of the observations, deviating from the actual period due to the effects of sampling. To check for periodicities appearing in only part of our data, for errors in the application of the ephemeris or the Solar System barycentre corrections, and for deviations of the Crab pulsar from the ephemeris, periodicity was searched for by sweeping around the pulsar's expected period. The counting rate of the array varies with time due to the changing zenith angle of the source, giving rise to a further spreading of the sampling range for a periodicity. Starting at 09.14 UT on 23 October, a strong effect was noticed at the directly adjacent Fourier harmonic to the ephemeris period (the harmonic separation being based on the total duration of the observation). The chance probability using the Rayleigh test was 10^{-4} with 30 d.f., which was felt to be an indicator of pulsed activity; to become consistent with the ephemeris period (that is, at the same Fourier harmonic), the duration of the pulsed emission was required to be shorter than the duration of the observation.

Each night's data were therefore split into 15 min sections, the length being dictated by that required to make the effect seen on 23 October occur in the same Fourier harmonic as the ephemeris period. The sections were each unfolded with the appropriate barycentric period, and the probability of uniformity calculated. Using this procedure the number of d.f. had been expanded from 1, in the case of the whole 34 h treated as one set, to 130 d.f. Nevertheless, the chance probability for one of the sections taken on 23 October decreased to $<10^{-6}$. On 31 October another section showed possible activity with a chance probability of $\sim 10^{-4}$.

To refine the positions of the apparent bursts of periodic emission and also to verify the derived probability values by observing the 'noise' spectrum, a section of fixed length moving in increments of ~ 2 min was used. The result of this procedure is a profile of a data set giving the probability of lack of periodicity (for 1 d.f.) as a function of time with a characteristic duration of 15 min. The profiles of the sets for two nights, 23 and 31 October, are shown in Fig. 1. The total number of d.f. has now expanded to $\sim 1,000$ because of the overlapping of the sections. The peak near the end of the run on 23 October has a Rayleigh test χ^2 value of 37.2 for 2 (internal) d.f., which has a chance probability of 4×10^{-9} . At this peak, $34 \pm 5\%$ of the recorded events are periodic, corresponding to an integral flux of $(2 \pm 0.3) \times 10^{-10} \text{ cm}^{-2} \text{ s}^{-1}$ for the 15 min burst. The threshold energy depends on the zenith angle at which the data are taken and for this observation was 3,000 GeV.

The records of showers arriving during the bursts have been examined for evidence of any properties different from those of 'normal' showers. The total number of showers is higher than that estimated from a period immediately following the burst on 23 October by $18\% \pm 7\%$, but the lack of gain stability procedures make the estimated number of background events imprecise. Our well separated (pitch = 60 m) array of four telescopes enables the showers to be examined in more detail. For example, those showers triggering two or more of the telescopes may be expected to arise preferentially from primary γ rays. This is because of the earlier development of γ -ray initiated showers in the atmosphere, resulting in a slightly wider light pool on the ground. During the burst on 23 October, the showers had been separated into two groups: those occurring during a 20% phase interval at the peak of the light curve, and the rest. The latter group was similar in all ways to the showers in the interval following the burst, while the former group was richer in double telescope responses, with a χ^2 chance probability of 1% which could indicate a wider light pool for some of the events. An alternative explanation could be that the on-phase events possess a harder energy spectrum but an analysis

of the photon densities recorded by the individual phototubes shows no significant difference between the two groups. The number of higher multiplicity responses during the 15 min burst was inadequate for this comparison, and the strength of the possible burst on 31 October was insufficient to permit a similar analysis.

The light curve of the Crab pulsar at much lower γ -ray energies has been reported by the Caravane collaboration. In early observations, it showed a main and an interpulse of a few milliseconds duration, and in later observations the interpulse was seen to decrease in amplitude⁹. The light curves inferred from our events occurring during the bursts on 23 and 31 October are shown in Fig. 2 with the relative phase provided by the ephemeris used. All of them show a single pulse of similar shape and long duration of 6 ms FWHM. It is not clear at what absolute phase, with respect to the radio, optical and X-ray pulses, the emission of UHE γ rays is to be expected, nor whether interpulses should occur at this vastly higher

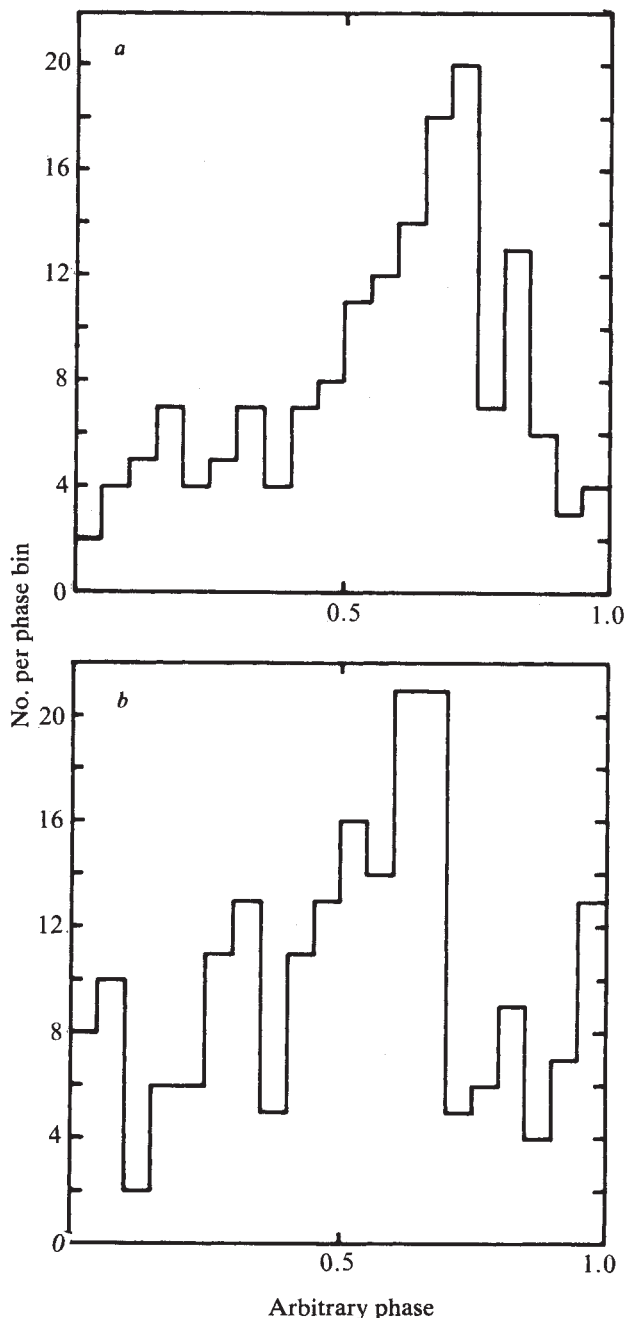


Fig. 2 The light curves for the periods of activity on a, 23 October 1981 and b, 31 October 1981.

photon energy. The derivation of the absolute phase of the light curves in Fig. 2 will require a more detailed ephemeris than that which is now available.

In conclusion, we have shown that the Crab pulsar may be emitting pulsed UHE γ rays in occasional short bursts. A previous balloon experiment¹⁰ at an energy of 1 GeV showed strong variability in pulsed flux between two flights, and during the flight providing the stronger flux, some evidence for variability on a time scale of 1 h. Earlier experiments had reported, at an energy threshold of 2,000 GeV, a time-averaged flux and upper limits near $10^{-11} \text{ cm}^{-2} \text{ s}^{-1}$. The flux of UHE γ rays in the bursts reported above, if averaged over 34 h of observation, suggests a similar time-averaged value and accounts for the previous lack of reproducibility in a straightforward way.

We thank the SERC for funding and for research studentships for I.W.K. and M.W.; Drs T.C. Weekes and H. Tornabene for the loan of equipment; and the commander and staff of Dugway Proving Ground, Utah for supporting our experimental work.

Received 1 February; accepted 11 March 1982.

1. Grindlay, J. E., Helmken, H. F. & Weekes, T. C. *Astrophys. J.* **209**, 592–601 (1976).
2. Jennings, D. M. *et al. Nuovo Cimento* **20B**, 71–82 (1974).
3. Erickson, R. A., Fickle, R. K. & Lamb, R. C. *Astrophys. J.* **210**, 539–543 (1976).
4. Gupta, S. K., Ramana Murthy, P. V., Sreekantan, B. V. & Tonwar, S. C. *Astrophys. J.* **221**, 268–273 (1978).
5. Sreekantan, B. V. *Phil. Trans. R. Soc. A* **301**, 629–632 (1981).
6. Gibson, A. I. *et al. Proc. 17th int. Conf. Cosmic Rays*, **8**, 38–41 (1981).
7. Mardia, K. V. *Statistics of Directional Data* (Academic, New York, 1972).
8. Bloomfield, P. *Fourier Analysis of Time Series* (Wiley, New York, 1976).
9. Wills, R. D. *Phil. Trans. R. Soc. A* **301**, 537–539 (1981).
10. Greisen, K. *et al. Astrophys. J.* **197**, 471–479 (1975).

Discontinuous precipitation reaction in the metal of Richardton chondrite

H. J. Axon & V. J. Grokhovsky

Metallurgy Department, University of Manchester, Grosvenor Street, Manchester M1 7HS, UK

A metallographic and microprobe study was conducted on the H5 chondritic meteorite Richardton, using material that was free of terrestrial corrosion and ablation heating effects. Our observations indicate that initially well annealed and slowly cooled metal experienced an incomplete episode of discontinuous precipitation in the temperature range 350–400 °C, perhaps after mild deformation at some of the metal–silicate interfaces. There were no indications of externally imposed deformation after the 350 °C episode. The mechanism of discontinuous precipitation rather than plane interface growth has implications for the rate of formation of the metallographic structures, and the apparent temperature of this final heat treatment has implications for proposed reheating events identified by isotopic measurements of the silicate minerals that accompany the metal in this meteorite.

The meteorite was observed to fall in 1918 and was originally catalogued¹ in terms of its macroscopic features as a veined spherical olivine–bronzite chondrite. In modern terminology² this corresponds to the H family of high iron content. The microstructural condition of its silicate minerals has been reported as appropriate to petrological type 5, which Dodd³ has equated with a metamorphic temperature in the range 700–800 °C, followed by slow (1°C My^{-1}) cooling to below 500 °C.

Affitalab and Wasson⁴ measured the Ni and Co content of α , kamacite and γ taenite, in Richardton. Their $\alpha:\gamma$ relationship crossed the tie-lines derived by Sears and Axon⁵ for un-reheated H chondrites and was thus more appropriate to the reheated material examined by Sears and Axon⁵. However, Affitalab and Wasson⁴ did not consider the possibility of reheating effects in their sample.

Turner *et al.*⁶ investigated the argon release from whole-rock samples of Richardton and derived a total argon release age of $4,500 \pm 30$ Myr. However, they noted that “significant small fluctuations in apparent age occurred over the extent of the release pattern”. This observation might suggest some sort of reheating episode.

By contrast, Evensen *et al.*⁷ obtained a Rb/Sr age of $4,390 \pm 30$ Myr from a study of separate chondrules which they removed from the meteorite.

Thus, from isotope studies of the silicates in Richardton there are indications of reheating additional to that encountered in other H chondrites, but so far there has been no special metallographic study of possible reheating effects as shown by the metal.

Metal in Richardton is usually present with troilite, FeS, and may be encountered either as very small (10 μm) particles embedded within the silicate chondrules, or, as larger (>100 μm) bodies, approaching chondrule size, in the ‘matrix’ between silicate chondrules.

Bevan and Axon⁸ have drawn attention to the existence of metallic chondrules as a feature of the pristine agglomeration structure of chondrites. Metallic chondrules may contain varying proportions of sulphide and silicate phases and thus they complement the spectrum of silicate/metal proportions and structures that have been encountered in the well-recognized silicate chondrules. However, metal is more responsive than silicate to metamorphic forces, thus the matrix metal may be subject to changes of shape by plastic deformation and to changes of shape and size and $\alpha:\gamma$ structure by diffusive mass transfer. Hence the chondritic nature of matrix metal may become obscured more rapidly than the silicate structures.

The present study was conducted on metal of the larger (=matrix) type. The samples were from unused residual material from the work of Evensen *et al.*⁷ (presented by O’Nions). They were mounted as whole-rock samples for metallographic, electron probe analysis and SEM investigation.

Figure 1 shows a typical area of metal as it appears after etching in Nital. The main features of this structure are a new generation of smaller transformation grains of T-kamacite (low Ni, α , b.c.c. Fe–Ni) growing inward from the metal–silicate

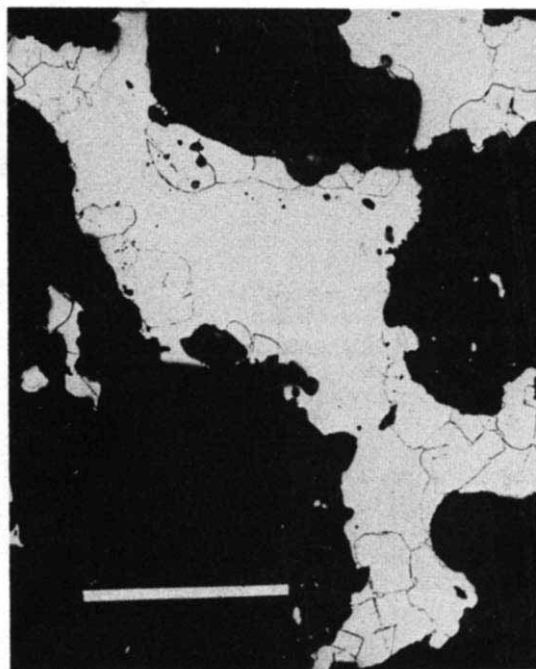
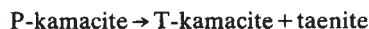


Fig. 1 The structure of a typical matrix metal area in the Richardton (H5) chondrite. Scale bar, 100 μm . At the metal–silicate edge of the large grain of parent kamacite, ~25 smaller grains of transformation structure have developed. There is a small-scale precipitation of γ taenite within the transformation structure; see Fig. 2.

boundary and into the pre-existing monocrystal mass of parent P-kamacite, see Fig. 2. Both P- and T-kamacite are remarkably free from Neumann bands.

Figure 2 shows that the growth of T has ceased and a parent-transformation interface is detectable. The transformation kamacite is decorated with small elongated particles of taenite (high Ni, γ , f.c.c. Fe-Ni) which have developed perpendicular to the parent-transformation (P-T) reaction front. This structure is essentially that of discontinuous precipitation of the type⁹



The location and distribution of the T grains suggests that they always nucleated at points of local strain on the silicate-metal interface. It appears that cold metal had been subjected to deformation at high-spots in the surrounding silicate, but it is not certain whether the deformation was induced by externally applied forces or was generated by differential contraction between metal and silicate. However, the stress-induced nucleation of the transformation product is confirmed by the occasional appearance of polygonized strain lines within the parent kamacite where it comes into contact with large, unbroken, silicate bodies.

Note that there are additional strain lines that appear to have been formed by the strain effects that accompanied or were generated by the actual progress of the reaction, but this strain is different from that in the nucleating process that we propose above. Some of these considerations have been reviewed elsewhere¹⁰ for the case of manmade alloys.

As usual with chondritic metal¹¹, there is some scatter of bulk nickel content between different individual grains of P-kamacite. However, within each single P+T-kamacite assemblage of Richardton there is a distinct change of nickel content

at the transformation interface. The microprobe trace of Ni content across the transformation interface is illustrated in Fig. 2. The small taenite precipitates usually give microprobe analyses in the range 49–50% Ni with a maximum of 52% Ni. The proportions of T-kamacite to precipitated taenite are visually estimated as sufficient to maintain a nickel mass-balance between the parent kamacite and the transformed regions. Reference to the Fe-Ni equilibrium diagram¹² indicates that a taenite of 50% nickel is the equilibrium precipitate at temperatures below $\sim 400^\circ\text{C}$.

Thus the composition and microstructure of the matrix metal in Richardton suggests that it developed a shock-free structure at a temperature appropriate to the metamorphic type 5 of its silicates—perhaps 700–800 $^\circ\text{C}$. Slow cooling to the temperature of maximum α solid solubility ($\sim 450^\circ\text{C}$) allowed the development of P-kamacite with a nickel content at least approaching that maximum.

Below 450°C cooling may have followed one of two routes, of which we prefer the second. (1) The decreased α solid solubility below 450°C brought the P-kamacite into a supersaturated condition from which small particles of taenite precipitated in a discontinuous mode as cooling continued. The cellular growth which is visible in Figs 1 and 2 leaves unresolved the mechanism whereby nucleation was induced if the cooling is assumed to have been smooth and continuous. (2) Alternatively, the supersaturated, monocrystal, P-kamacite may have been mildly stressed at locations where relatively large and unbroken bodies of hard silicate made contact with and were able to deform the metal. The temperature of this event cannot be specified but must have been in the low temperature, cold working, range. The deformation was of a slow and gentle nature. The absence of Neumann bands indicated that it was not produced by shock. The evidence lies in the local development of transformation where kamacite is in contact with coarse unbroken silicate bodies but not where it abuts fine matrix material and in the occasional presence of un-recrystallized strain markings around penetrative silicate bodies in the parent kamacite.

Reheating to 350–400 $^\circ\text{C}$ after strain at a lower temperature gives a more unstable condition in the kamacite than is obtained on continuous cooling to that temperature, hence our preference for this production route. In the case of reheating, the taenite precipitates are to be contrasted with the isothermal taenite produced¹³ at higher temperatures in reheated Cañon Diablo kamacite.

The discontinuous precipitation process, now identified for the first time in the cosmic metal of Richardton, involves diffusion within and parallel to the moving transformation front, whereas the previously recognized production of clear or zoned taenite involves diffusion perpendicular to the growth interface. The time required for the production in Richardton of the discontinuous precipitation structure is probably less than would be required for plane interface growth but, nevertheless, appears to be considerable and, if more detailed quantitative information were available an estimate might be attempted. Finally, there have been several studies of meteoritic material artificially reheated at 500 $^\circ\text{C}$ and above^{13,14} but very little investigation at lower temperatures. To study this possibility we have examined samples of Cañon Diablo annealed for 4 weeks at 450 $^\circ\text{C}$ and 400 $^\circ\text{C}$. At 400 $^\circ\text{C}$ Neumann band alteration was detectable within the mass of the sample but, in addition, at 450 $^\circ\text{C}$ there was surface recrystallization, similar to that in Fig. 1 but on a smaller scale, at the cut, and hence deformed edges of the annealed sample.

Received 14 December 1981; accepted 15 March 1982.

1. Hey, M. H. *Catalogue of Meteorites* (British Museum, London, 1966).
2. Van Schmus, W. R. & Wood, J. A. *Geochim. cosmochim. Acta* **31**, 747 (1967).
3. Dodd, R. T. *Geochim. cosmochim. Acta* **33**, 161 (1969).
4. Affialalab, F. & Wasson, J. *Geochim. cosmochim. Acta*, **44**, 431 (1980).
5. Sears, D. W. & Axon, H. J. *Nature* **260**, 34 (1976).
6. Turner, G., Enright, M. C. & Cadogan, P. H. *Proc. 9th Lunar planet. Sci. Conf.* 989 (1978).
7. Evensen, N. M., Carter, S. R., Hamilton, P. J., O'Nions, R. K. & Ridley, W. I. *Earth planet. Sci. Lett.* **42**, 223 (1979).

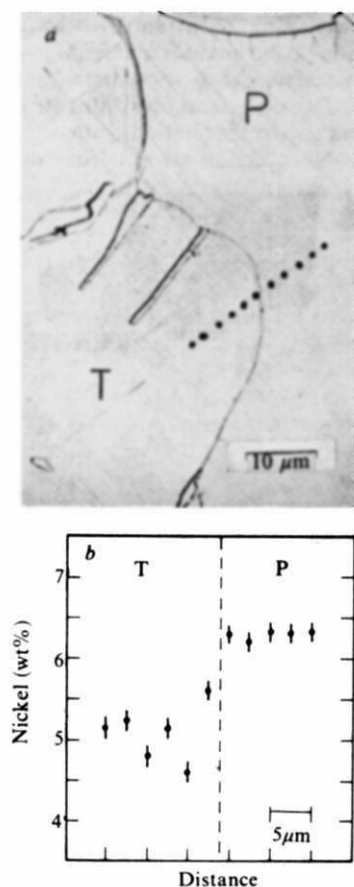


Fig. 2 An enlarged view (a) of the interface between P-(parent) and T-(transformed) kamacite, with the location of microprobe analyses and corresponding (b) Ni contents across the P-T interface. The elongated precipitates perpendicular to the P-T interface are γ taenite containing $\sim 50\%$ Ni.

8. Bevan, A. W. R. & Axon, H. J. *Earth planet. Sci. Lett.* **47**, 353 (1980).
9. Gust, W. *Phase Transformations Ser. 3*, No. 11, Ch. 1, 11 (The Institution of Metallurgists, London, 1979).
10. Williams, D. B. & Butler, E. P. *Int. Metal. Rev.* No. 3, 153 (1981).
11. Hutchison, R., Bevan, A. W. R., Easton, A. J. & Agrell, S. O. *Proc. R. Soc. A* **374**, 159 (1981).
12. Romig, A. D. & Goldstein, J. I. *Geochim. cosmochim. Acta* **45**, 1187 (1981).
13. Brentnall, W. D. & Axon, H. J. *J. Iron Steel Inst.* **200**, 947 (1962).
14. Buchwald, V. F. *Handbook of Iron Meteorites Vol. 1*, 131–136 (University of California Press, Berkeley, 1975).

Distribution and enantiomeric composition of amino acids in the Murchison meteorite

Michael H. Engel* & Bartholomew Nagy

Laboratory of Organic Geochemistry, Department of Geosciences, The University of Arizona, Tucson, Arizona 85721, USA

Early determinations of the amino acid distribution in water extracts of the Murchison meteorite revealed unusual amino acids, including isovaline (Ival), α -aminoisobutyric acid (α -Aiba) and pseudoleucine (Ple), as well as common ones such as glycine (Gly), alanine (Ala) and glutamic acid (Glu)^{1–10}. Amino acids that could be resolved into their respective D- and L-enantiomers were reported to be racemic (see, for example, refs 1, 3, 5), although small sample size might have hindered the precise determination of D/L values. Using improved chromatographic and mass spectrometric procedures we have now been able to amplify and resolve the partially racemized amino acids, Glu, aspartic acid (Asp), proline (Pro), leucine (Leu) and Ala, in an interior sample of a Murchison meteorite stone. Water-extractable amino acids were more racemized than those recovered by digesting the water-extracted meteorite in 6M HCl. The amino acid composition of this stone was similar to previous reports^{1–8}, including the absence of tyrosine (Tyr), methionine (Met), phenylalanine (Phe) and only minor traces of serine (Ser) and threonine (Thr)⁹. Serine and Thr are usually considered to be terrestrial contaminations. This is the first report of amino acids in a carbonaceous meteorite which, based on currently accepted criteria, appear to be indigenous but for unknown reasons are not racemic. Confirmation of these findings by other investigators suggests that further examination of amino acids in clean carbonaceous meteorites could be beneficial.

We have analysed a single Murchison meteorite stone (Me 2640, Field Museum of Natural History, Chicago, Illinois; 50.7 g) which has over 90% of its surface covered with an intact fusion crust. Standard precautions were taken to prevent laboratory contaminations^{11–13}. First, the entire surface (~2 g) of the stone was removed with a high-speed drill. Several interior fragments (39.0 g) were ground to a fine powder using a porcelain mortar and pestle. The meteorite powder and a procedure blank were water-refluxed for 8 h at 100 °C. The water extracts were hydrolysed in 6M HCl for 24 h at 100 °C. The procedure used for cation exchange and derivatization of the water-extractable amino acids has been reported previously¹³. To determine whether hydroxy amino acids, such as Ser and Tyr, were present, the *N*-PFP, *O*-PFP-(+)-2-butyl esters were converted to *N*-PFP, *O*-acetyl-(+)-2-butyl esters^{14,15} for analysis on effective 61-m long 0.5 mm internal diameter nickel capillary columns coated with Carbowax-20M. Next, the water-extracted residue was dried and then digested in 6M HCl for 24 h at 100 °C. In a control experiment¹³, HCl digestion did not racemize or destroy amino acids in the presence of iron oxides and silicate minerals. The Murchison acid extract was desalted using an anion exchange column followed

by a cation exchange column^{15,16}. The amino acids were then derivatized to *N*-PFP, *O*-PFP-(+)-2-butyl esters.

Gas chromatograms of the Murchison water and 6M HCl extracts and of the procedure blank are shown in Fig. 1. Amino acid identifications were confirmed by retention times and co-injection of amino acid standards. The hydrolysed water extract was re-analysed by combined gas chromatography–chemical ionization mass spectrometry (GC–CIMS). Amino acid identifications were confirmed by comparing unknown mass spectra with mass spectra of derivatized amino acid standards (see, for example, Fig. 2). Amino acids recovered by digesting the water-extracted meteorite powder in 6M HCl were, as previously reported¹, less abundant than those removed by water extraction, but the samples were too small to confirm their identities by GC–CIMS. Instead they were identified by co-injection and retention times of amino acid standards whose identifications were confirmed by GC–CIMS.

Amino acid D/L values are listed in Table 1. Some amino acids, such as L-Asp and γ -amino-*n*-butyric acid (GABA), co-eluted; a small, unknown peak co-eluted with D-Pro. For D/L values of carbonaceous meteorites to be significant, the purity of individual amino acid enantiomers must be confirmed by a technique that can subtract co-eluting compounds. Selective ion monitoring (ref. 17 for example) made it possible to remove interfering GC peaks and to obtain and confirm D/L values for Asp, Pro, Glu and Ala in the hydrolysed Murchison water extract. For example, three ions characteristic to Asp (*m/e* = 420, 280, 262) in CIMS were not present in the overlapping GABA mass spectra. The D/L values for selected ion traces were determined from ion current peak heights. Good correlations were found between D/L values of Ala and Glu determined from integrated GC peak areas and from relative ion intensities. As predicted, the D/L value for Asp was slightly higher (0.30) and that for Pro slightly lower (0.30) when peak interferences were eliminated (Table 1). Figure 3 shows traces of intensities of the three ions common to Ala and not present in overlapping or adjacent amino acid GC peaks of the Murchison water extract.

In these experiments, most of the water-extractable amino acids were partially racemized, whereas the 6M HCl-extractable amino acids were less racemized. These results may not contradict previous reports, because we have analysed a large interior Murchison sample (39.0 g), ensuring recovery of an abundant amino acid fraction for enhanced gas chromatograms as well as for the GC–CIMS ion traces.

The abundances of amino acids detected in the hydrolysed Murchison water and 6M HCl extracts are listed in Table 2. We detected none of the common protein amino acids such as Ser, Thr, Tyr, Phe and Met, that are usually absent in carbonaceous meteorites, in these extracts, although minor traces of Ser and Thr were detected by Cronin and Pizzarello (personal communication) in this stone. Serine is the, or one of the, major components of fingerprints, whereas Thr, Tyr, Met, Phe, Glu, Asp, Pro, Leu and Ala are much less abundant. This suggests that originally the stone did not contain Ser, Thr, Tyr, Phe and Met, and that the other amino acids are not fingerprint contaminations. The four most abundant amino acids in the hydrolysed water extract of the Murchison meteorite were α -Aiba, Gly, Glu and Ala. Cronin *et al.* have reported similar findings⁹.

The absence of Ser, Tyr, Phe and Met in the Murchison meteorite extracts might have been due to their decomposition during sample extraction. To investigate this possibility, we prepared duplicate mixtures consisting of ~0.3 mg of Phe, Tyr, Ser, Met, Glu and Ival. Two hundred milligrammes of the unextracted Murchison meteorite powder were added to one of the mixtures; the other mixture served as a control. Water (1 ml) was added to both mixtures, which were extracted and then analysed for their amino acid contents using the procedure described above for the analysis of the Murchison meteorite water extract. The quantities of Phe (52%), Tyr (13%), Ser (45%) and Met (35%) lost from the mixture containing the meteorite powder during extraction and/or cation exchange

* Present address: Geophysical Laboratory, Carnegie Institution of Washington, 2801 Upton St North-west, Washington, DC 20008, USA.

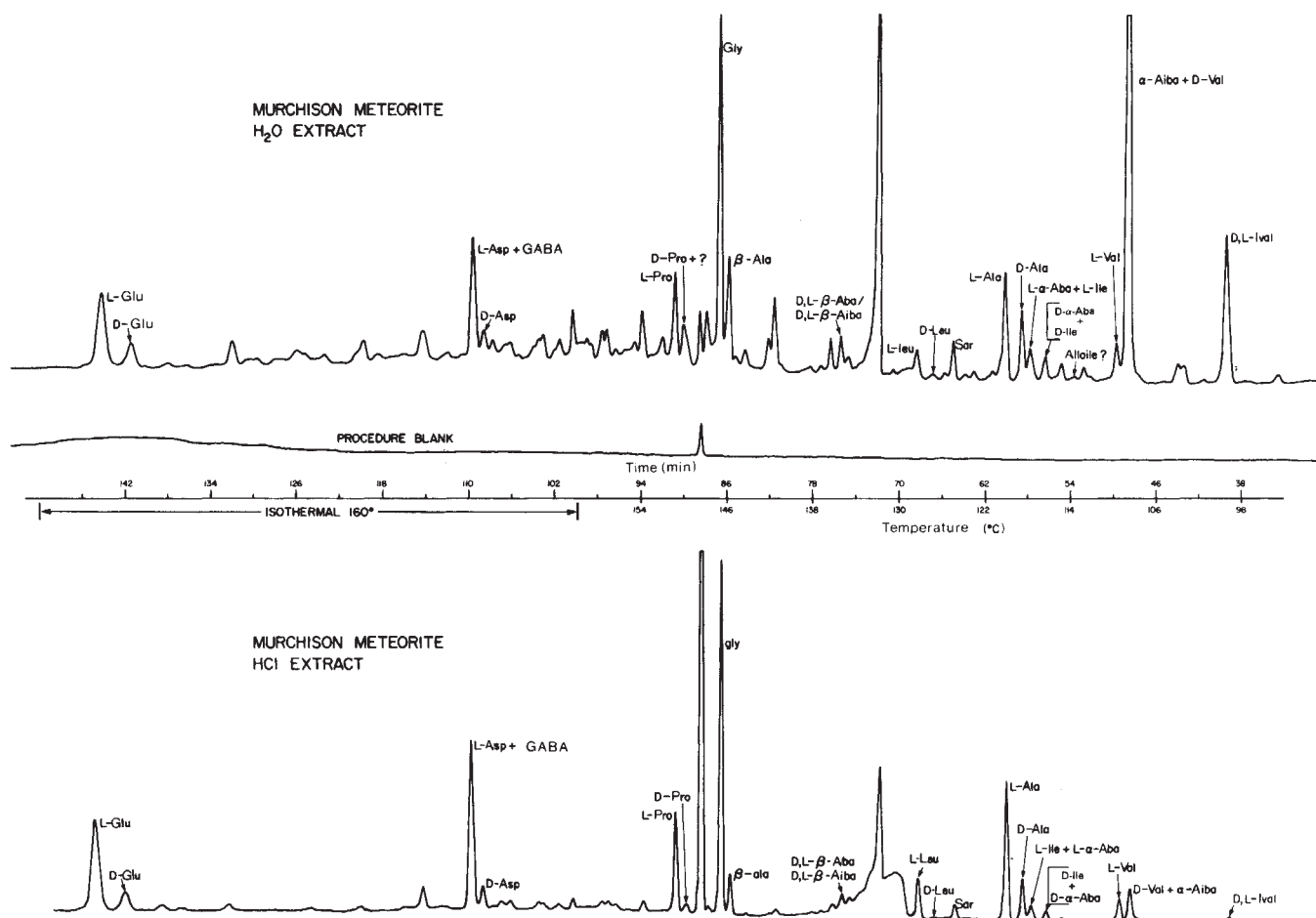


Fig. 1 Gas chromatograms of the procedure blank, water and 6M HCl extracts of the Murchison meteorite. In addition to co-injection of amino acid standards and retention times, all amino acid identifications for the Murchison water extract were confirmed by combined GC-CIMS. The GC and GC-CIMS conditions were identical to those previously reported¹³.

were similar to those detected for Glu (40%) and Ival (45%), indicating that if Phe, Tyr, Ser and Met were present in this meteorite stone, they would probably have been detected. Also, there was no deviation from the original D/L values in the control experiments.

The GC and GC-CIMS techniques did not resolve D- and L-Ival. The co-elution of D- and L- α -amino-*n*-butyric acid (α -Aba) with D- and L-isoleucine (Ile) and their low concentrations made it impractical to use selective ion monitoring to determine their D/L values. Cronin *et al.*¹⁸ have reported that some amino acid constituents of carbonaceous meteorites, such as Ival and α -Aba, are susceptible to decomposition reactions, which appear to correlate directly with increased residence time of the meteorite on Earth. To resolve the respective D- and L-enantiomers of Ival and α -Aba from the Murchison meteorite, a 3.0 g sample of the Murchison meteorite powder was refluxed twice in water for 2 h at 80 °C. The samples were centrifuged, and the extracts (I and II) were evaporated to dryness and cation exchanged, and then redissolved in water. A slightly modified HPLC procedure¹⁹ was used to isolate and resolve the D- and L-enantiomers of Ival and α -Aba from the two extracts. Isovaline, as previously reported²⁰, was racemic in extracts I (D/L = 0.99 ± 0.06) and II (D/L = 0.99 ± 0.04). α -Aba appeared to be racemic (extract I, D/L = ~ 1.0 ; extract II, D/L = 1.13 ± 0.23) although its low concentration in the extracts made precise determination of D/L values impossible (Fig. 4).

Some of our results have been independently confirmed in other laboratories. Amino acid abundances in a hot water extract (24 h, 110 °C) of part of the same Murchison meteorite stone (Me 2640) examined by ion-exchange analysis with fluorometric detection (J. R. Cronin and S. Pizzarello, personal communication) were reasonably consistent with those in Table

2. Sample Me 2640 seemed to contain amino acids typical of the Murchison carbonaceous meteorite stones. Fingerprint amino acids, such as Ser, were in low concentrations, and Tyr, Met and Phe were absent. The amino acid abundances and distribution pattern seem virtually to rule out contamination from handling or microorganisms. It is interesting that this stone was enriched in α -substituted amino acids, for example Ival, 2-amino-2,3-dimethylbutyric acid and 2-methylnorvaline, and had fewer amino acids with an α -hydrogen atom than another Murchison stone⁹.

Table 1 Murchison meteorite amino acid D/L values

Extract	Glu	Asp	Pro	Leu	Ala
H ₂ O*	0.322 (± 0.044)	0.202 (± 0.005)	0.342 $\pm$ (± 0.065)	0.166 $\pm$ (± 0.021)	0.682 $\pm$ (± 0.062)
H ₂ O†	0.30 (± 0.02)	0.30 (± 0.04)	0.30 (± 0.02)	ND	0.60 (± 0.03)
HCl*	0.176 (± 0.013)	0.126 (± 0.004)	0.105 (± 0.017)	0.029 (± 0.002)	0.307 (± 0.010)

The experimental errors for the D/L values are shown in parentheses. Approximately 2.0% of the D-amino acids used to obtain the D/L values may be attributed to (γ -2)-butanol impurities and racemization during 6M HCl hydrolysis. ND, poor resolution did not permit calculation of this value by selected ion traces.

* D/L values were calculated from integrated peak areas and are an average of at least three gas chromatographic determinations.

† D/L values were calculated from selected ion peak heights and are an average of D/L values determined from three specific ions monitored for each amino acid.

‡ One or more D/L values determined by weighing the peaks (in addition to digital integration).

Table 2 Murchison meteorite amino acid concentrations

Extract	Glu	Asp	GABA	Pro	Gly	β -Ala	β -Aiba nmol g ⁻¹	Leu	Sar	Ala	Ile α -Aba	Val	α -Aiba	Ival
H ₂ O	18.15	8.48	6.28	13.45	45.79	13.07	5.22	1.91	4.69	15.34	5.98	8.59	107.84	23.57
HCl	6.80	4.02*	3.22*	3.35	14.12	1.25	0.96	0.88	0.57	4.84	0.89	1.09*	1.82*	0.20

D,L- β -Aba and D,L- β -Aiba and D,L-Ile and D,L- α -Aba co-eluted and could not be resolved by GC or selected ion currents because of low abundances and the similarity of mass spectra, respectively. Their respective abundances are reported for the Murchison extracts as average contributions of the two.

* Concentrations were estimated from the relative abundances determined for the Murchison water extract by selected ion monitoring.

A portion of this Murchison meteorite stone and a blank were extracted with hot water and derivatized (*N*-PFP, *O*-PFP-(+)-2-butyl esters) by Cronin and Pizzarello (personal communication), and then sent to P. E. Hare for independent GC determinations of the amino acid D/L values. A Perkin-Elmer 900 gas chromatograph with a Perkin-Elmer nitrogen detector, equipped with a 50.0 m $\times$ 0.3 mm i.d. glass W.C.O.T. chiralasil capillary column, was used. The blank contained only one very small peak with a retention time close to Gly. The D/L values in the derivatized hot water extract of the meteorite were D/L Ala = 0.79, D/L Pro = 0.65 and D/L Glu = 0.56 (P. E. Hare, personal communication).

Only 24 naturally occurring peptides and proteins (3–31 residues long) have been reported not to contain Ser, Tyr, Phe and Met. All are very uncommon in organisms; for example antifreeze glycoprotein 3 (31 residues, isolated from the fish

Trematomus borchgrevinki), fibrinopeptide B (10 residues, isolated from rhinoceros) and thyroliberin (3 residues, isolated from pig) (M. O. Dayhoff, personal communication). If the Murchison stone we analysed was more than minimally contaminated, it would probably have contained significant quantities of Ser, Tyr, Phe and Met.

The presence of L-Ser in carbonaceous meteorites is generally considered to represent a terrestrial contribution. Assuming that small quantities of L-Glu, L-Ala and L-Pro were introduced into the Murchison stone at the same low level as Ser (0.9 nmol g⁻¹; Cronin and Pizzarello, personal communication), the D/L values can be recalculated after decreasing the L-amino acid abundances proportional to the trace amount of Ser found in this sample. This correction, however, does not significantly alter the D/L values.

Several hypotheses may at least partially account for the occurrence of non-racemic amino acids in this Murchison meteorite stone: (1) terrestrial contamination during entry, impact, collection and/or in the laboratory; (2) various unknown preferential syntheses of non-racemic amino acids during laboratory procedures; and (3) unknown stereoselective chemical synthesis and/or decomposition reactions unrelated to terrestrial sources.

(1) If the stone was more than minimally contaminated before or during impact and/or at any time thereafter, the absence of Tyr, Phe, Met and only a trace of Ser, would not be predicted, as these amino acids are ubiquitous to terrestrial biogenic systems. Also, the $\delta^{15}\text{N}$ values of amino acids and other nitrogenous compounds in Murchison are very different from terrestrial $\delta^{15}\text{N}$ values²¹. If terrestrial contamination by L-amino acids is invoked to account for the Murchison meteorite D/L values, there is either a very unusual contamination source of specific amino acids, or preferential adsorption and/or decomposition of specific amino acids, by interactions at probably ambient temperatures of the meteorite organic/inorganic matrices with the intruding material(s). Although preferential adsorption of L-amino acids by mineral matrices has been reported (ref. 22 for example), the adsorption or destruction of L-amino acids such as Ala, Glu, Asp, Pro and Leu in preference to others, for example Ser, Tyr, Phe and Met has not. Amino acids extracted from the meteorite with 6M HCl were less racemized than the water-extractable amino acids. To account for these ratios by contamination, the terrestrial contaminants would have to be preferentially and very tightly complexed to the meteorite matrix. In previous experiments in which rocks were soaked in aqueous amino acid solutions for 1,321 h, all the amino acids entering the samples during soaking were extracted from the rocks using the water extraction procedure we used for the Murchison meteorite. Subsequent acid extraction resulted in no additional recovery of amino acids inserted into the rock samples (M. H. Engel and B. Nagy, in preparation). If the Murchison meteorite was contaminated with L-amino acids they should have been removed during the water extraction. L-amino acids introduced into the stone on Earth during the 11 yr before analysis would be predicted to be less tightly bound to, and more readily extracted with, water from the meteorite organic/inorganic matrix than amino acids incorporated in the meteorite some 4.5×10^9 yr ago. Yet, 6M HCl extraction recovered amino acids showing even less racemization than the water extract, so, in fact, the less racemic

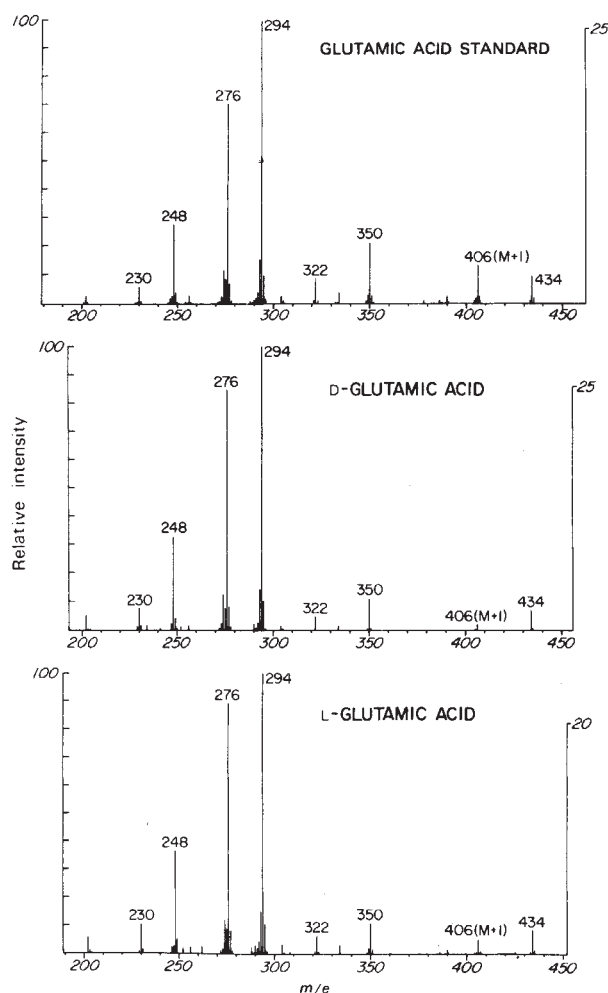


Fig. 2 Mass spectra of glutamic acid standard (top), and of D-glutamic acid (middle) and L-glutamic acid (bottom) in the water extract of the Murchison meteorite. GC-CIMS conditions are given in ref. 13.

amino acids are more tightly held in this stone than those extractable with water. The finding of non-racemic amino acids, however, makes contamination of the Murchison meteorite a serious possibility.

(2) Hydrogen cyanide polymers²³ and *N*-carbamyl amino acids^{7,8} have been proposed as amino acid precursors in the water extracts of carbonaceous meteorites. Hydrolysis of these precursors to free amino acids would not involve reactions at the asymmetric centres of these compounds, indicating, however, that such precursors, if they were present in this stone, were themselves non-racemic.

(3) Mechanisms have been proposed for the stereoselective decomposition of amino acids²⁴⁻²⁷. If the amino acids in the Murchison meteorite were initially racemic, and if these mechanisms prove applicable, they might indeed account for the D/L values observed. A few mechanisms (for example ref. 28) have been proposed for the extraterrestrial stereoselective synthesis of organic compounds, but it is highly unlikely that amino acids in the Murchison meteorite would be completely racemic had they initially been present in the L or D form²⁹.

We therefore conclude that the partially racemic amino acids in this Murchison meteorite stone are probably due to extraterrestrial stereoselective synthesis or decomposition reactions, although the possibility of unusual terrestrial contamination cannot be excluded. Different stones of the Murchison meteorite may be heterogeneous with respect to their carbon compound compositions.

Optical activity in carbonaceous meteorites infer the possibility of asymmetric chemical syntheses unrelated to terrestrial sources. Although we do not know the specific cause(s) of the optical activity, the independent confirmations of our results strengthen our conclusion that the Murchison meteorite sample we studied was clean and that some of the amino acids are partially racemic.

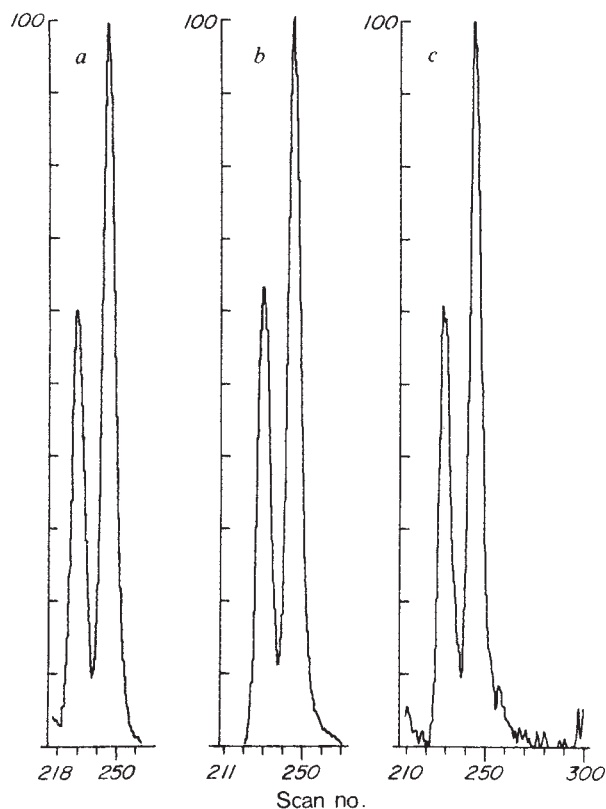


Fig. 3 Selected ion traces for alanine in the water extract of the Murchison meteorite. *a*, $m/e = 218$, D/L Ala = 0.60; *b*, $m/e = 216$, D/L Ala = 0.63; *c*, $m/e = 320$, D/L Ala = 0.58. $m/e = 320$ is the characteristic $M+29$ adduct formed by the addition of $C_2H_5^+$ to the alanine derivative during chemical ionization.

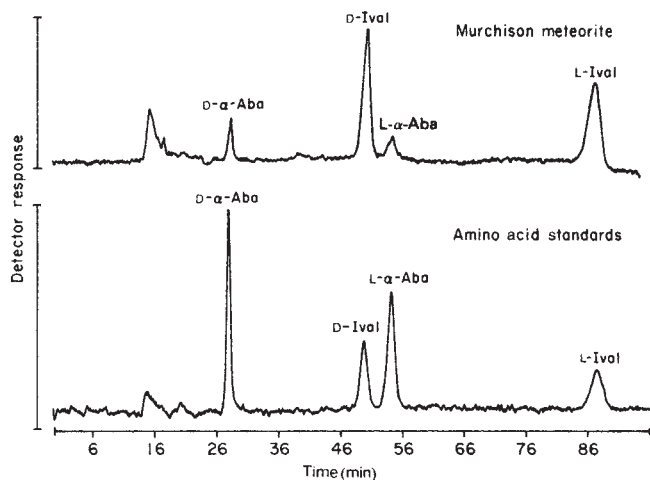


Fig. 4 Chromatograms of the D- and L-enantiomers of isovaline and α -amino-*n*-butyric acid from extract II of the Murchison meteorite (top) and of isovaline and α -amino-*n*-butyric acid standards (bottom). D, L-isovaline and α -amino-*n*-butyric acid were initially isolated from extract II by HPLC using a 25 cm $\times$ 2.0 mm i.d. stainless-steel cation exchange column packed with a 5 μ m cation exchange resin that had been converted to the pyridine form. An aqueous buffer consisting of pyridine (0.2 M) and formic acid (pH 3.2) was used. The buffer flow rate was 0.12 ml min⁻¹. The column temperature was 53 °C. The amino acids were collected using a stream splitter before reaction with *O*-phthalaldehyde and subsequent fluorometric detection. The D- and L-enantiomers of isovaline and α -amino-*n*-butyric acid were resolved using the procedure developed by Engel and Hare¹⁹ for the resolution of basic amino acid enantiomers.

We thank H. Ogino for helping to determine the relative stabilities of amino acids during sample extraction and to prepare additional Murchison meteorite extracts for analysis by HPLC, S. Y. Chang for help with the GC-CIMS analyses, J. R. Cronin, S. Pizzarello and P. E. Hare for their independent examinations of the meteorite sample, K. A. Kvenvolden, L. Margulis and S. L. Miller for comments and E. Olsen for providing the meteorite sample. This research was supported by NASA grant NGR 03-002-171.

Received 16 November 1981; accepted 26 February 1982.

- Kvenvolden, K. A. *et al.* *Nature* **228**, 923-926 (1970).
- Cronin, J. R. & Moore, C. B. *Science* **172**, 1327-1329 (1971).
- Kvenvolden, K. A., Lawless, J. G. & Ponnamperna, C. *Proc. natn. Acad. Sci. U.S.A.* **68**, 486-490 (1971).
- Oró, J., Gilbert, J., Lichtenstein, S., Wikström, S. & Flory, D. A. *Nature* **230**, 105-106 (1971).
- Lawless, J. G. *Geochim. cosmochim. Acta* **37**, 2207-2212 (1973).
- Cronin, J. R. & Moore, C. B. *Meteoritics* **9**, 327-328 (1974).
- Cronin, J. R. *Origins of Life* **7**, 337-342 (1976).
- Cronin, J. R. *Origins of Life* **7**, 343-348 (1976).
- Cronin, J. R., Pizzarello, S. & Moore, C. B. *Science* **206**, 335-337 (1979).
- Cronin, J. R., Gandy, W. E. & Pizzarello, S. *J. molec. Evol.* **17**, 265-272 (1981).
- Engel, M. H., Zumberge, J. E. & Nagy, B. *Analyt. Biochem.* **82**, 415-422 (1977).
- Zumberge, J. E., Engel, M. H. & Nagy, B. in *The Biogeochemistry of Amino Acids* (eds Hare, P. E., Hoering, T. C. & King, K. Jr) 503-525 (Wiley, New York, 1980).
- Nagy, B., Engel, M. H., Zumberge, J. E., Ogino, H. & Chang, S. *Nature* **289**, 53-56 (1981).
- Pollock, G. E. & Kawachi, A. H. *Analyt. Chem.* **40**, 1356-1358 (1968).
- Pollock, G. E., Cheng, C.-N. & Cronin, S. E. *Analyt. Chem.* **49**, 2-7 (1977).
- Pollock, G. E. & Kvenvolden, K. A. *Geochim. cosmochim. Acta* **42**, 1903-1905 (1978).
- Smith, G. G. & Wonnacott, D. M. in *The Biogeochemistry of Amino Acids* (eds Hare, P. E., Hoering, T. C. & King, K. Jr) 203-214 (Wiley, New York, 1980).
- Cronin, J. R., Moore, C. B. & Pizzarello, S. *Meteoritics* **15**, 277-278 (1980).
- Engel, M. H. & Hare, P. E. *Yb. Carnegie Instn Wash.* **80**, 394-397 (1981).
- Pollock, G. E., Cheng, C.-N., Cronin, S. E. & Kvenvolden, K. A. *Geochim. cosmochim. Acta* **39**, 1571-1573 (1975).
- Becker, R. H. & Epstein, S. *Geochim. cosmochim. Acta* **46**, 97-103 (1982).
- Bondy, S. C. & Harrington, M. E. *Science* **203**, 1243-1244 (1979).
- Matthews, C. N., Nelson, J., Varma, P. & Minard, R. *Science* **198**, 622-625 (1977).
- Bonner, W. A. & Flores, J. J. *Origins of Life* **6**, 187-194 (1975).
- Bonner, W. A., Van Dort, M. A. & Yearian, M. R. *Nature* **258**, 419-421 (1975).
- Noyes, P. H., Bonner, W. A. & Tomlin, J. A. *Origins of Life* **8**, 21-23 (1977).
- Keszthelyi, L. *Origins of Life* **7**, 349-354 (1976).
- Bloch, M. R. & Wirth, H. L. *Naturwissenschaften* **67**, 562-564 (1980).
- Schroeder, R. A. & Bada, J. L. *Earth Sci. Rev.* **12**, 347-391 (1976).

Structural features of todorokite intergrowths in manganese nodules

Shirley Turner*, Malcolm D. Siegel†
& Peter R. Buseck‡

* Department of Geology, and † Departments of Chemistry and Geology, Arizona State University, Tempe, Arizona 85287, USA

‡ Department of Geological Sciences, Harvard University, Cambridge, Massachusetts 02138, USA

The disordered, cryptocrystalline phases that comprise manganese nodules^{1,2} typically yield broad, diffuse X-ray lines that have led to ambiguity in their identification. Of particular interest is the manganese oxide phase that has been correlated to the presence of copper and nickel in nodules. This host phase has a characteristic X-ray *d*-spacing of ~ 10 Å and has been compared with both a naturally occurring material—todorokite^{3,4}—and a synthetic phase—buserite^{5,6}. The structures of those phases have been difficult to decipher, but recent work using high-resolution transmission electron microscopy (HRTEM) has provided insight into the todorokite⁷ and buserite⁸ structures. Using HRTEM results we now show that the nodule 10 Å phase does have the same structure as non-marine todorokite⁹. In addition, structural imaging shows that both nodule and non-marine todorokite exhibit similar disordering phenomena. A unique feature of the nodule todorokite is its intimate and topotactic intergrowth with a possible double-chain oxide. Its paragenetic relationship to todorokite is not yet determined.

By structural imaging of non-marine samples⁷, todorokite was shown to be one of three related intergrowth families that occur in manganese oxides. Each family consists of an intergrowth of variously sized tunnels that differ in sharing either single, double or triple chains of edge-shared $[\text{MnO}_6]$ octahedra. Todorokite structures share triple chains, and the large resultant tunnel structures allow for the inclusion of water and large cations. The proposed atomic arrangement of a common todorokite structure is shown in Fig. 1. This tunnel structure

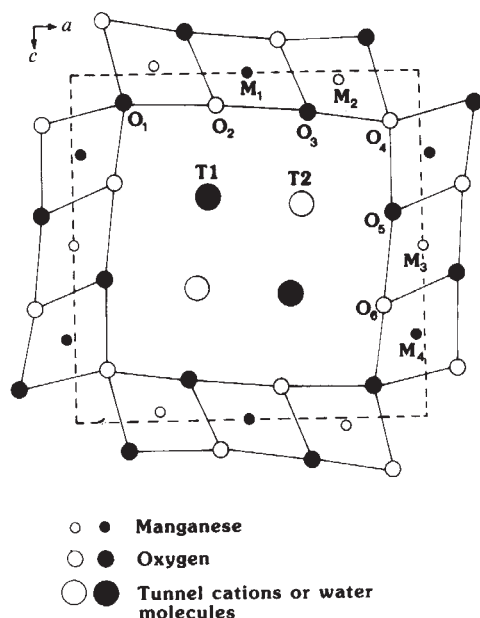


Fig. 1 The proposed atomic arrangement of a common todorokite structure, with three chains of manganese octahedra per side. The structure is projected down the tunnel direction $[010]$. $\circ$, $b = 0$; $\bullet$, $b = 1/2$. Dashed line denotes the unit cell.

leads to a fibrous habit, and these fibres are commonly twinned at 120° angles.

Presumed todorokite from manganese nodules has been studied previously by scanning and transmission electron microscopy. Both Siegel¹⁰ and Chukhrov *et al.*^{11–13} found fibrous material in nodules from the Pacific Ocean that contained twinning at the characteristic 120° angle; todorokites with different predominant tunnel sizes were found in these studies. In addition, Chukhrov *et al.*¹¹ show a high-resolution image of their nodule todorokite viewed perpendicular to the tunnel direction in which different tunnel sizes as well as disordering phenomena are evident.

The nodules we examined were collected from the North Equatorial Pacific near Manganese Nodule Project (MANOP) site S, which is between the Clarion and Clipperton Fracture Zones at about 140° W and 11° N. Details of the morphological and chemical features found in the nodules are presented elsewhere¹⁴. The nodule cores were found to consist largely of moulds of biogenic material such as discoasters, coccoliths and radiolaria. Scanning electron microscope studies of the nodules show that fibres of crystalline manganese oxide occur almost exclusively with the biogenic debris, an association that may have important paragenetic significance. Figure 2 shows

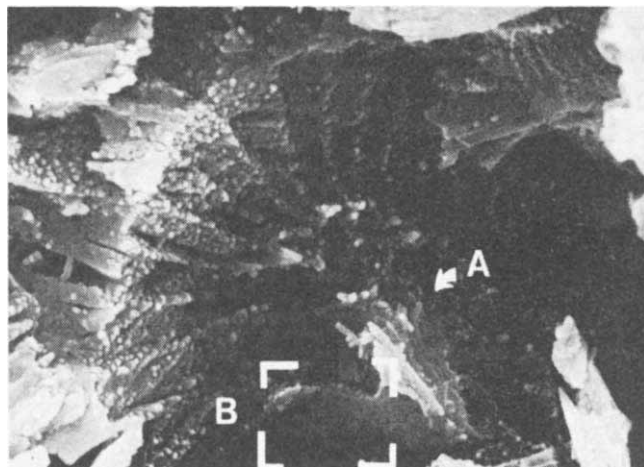


Fig. 2 Close-up of an impression of a coccolith shell. In the centre foreground, fibre tips of todorokite are evident (A). Todorokite fibres overlapping at about 120° angles can be seen in the box B.

examples of overlapping fibres and fibre tips associated with a mould of a coccolith shell. This fibrous morphology is similar to that of non-marine todorokite.

Samples of this core material were crushed and examined in a slightly modified JEOL JEM 100B microscope¹⁵. Due to their fibrous morphology, almost all grains settled onto the microscope grids with their lengths perpendicular to the electron beam direction. A structure image taken in this direction in Fig. 3a reveals triple, quadruple and quintuple chain widths. In other images, double chains, as well as sextuple and septuple chains, are shown. A view down the tunnel direction is necessary to determine the common shared width of these differently-sized tunnels. Unfortunately, the small cross-section of the nodule fibres precludes the use of electron diffraction for this orientation process. In one case, a fibre was fortuitously oriented nearly parallel to the electron beam. Figure 3b shows the image obtained, in which a slightly skewed framework of differently sized tunnels is evident. The estimated tunnel dimensions are labelled in the sketch of Fig. 3c. The image in Fig. 3b confirms that this nodule, 10 Å phase, is formed of an intergrowth of tunnel structures that have a triple chain of manganese octahedra in common. This fibrous 10 Å phase does then indeed have the same structure as non-marine todorokite.

Disordering phenomena are evident at all levels of

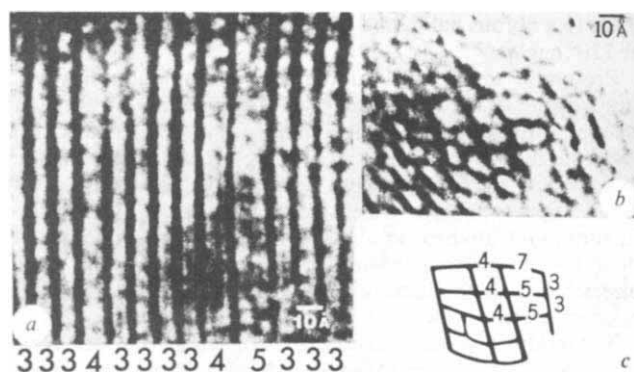


Fig. 3 *a*, HRTEM micrograph of todorokite from a nodule core, viewed perpendicular to the tunnel direction. The tunnels run from top to bottom in this image. Triple chains of manganese octahedra predominate, but quadruple and quintuple chains are also present. Note the bending of the tunnels on the right side. *b*, A nodule todorokite fibre oriented with its length parallel to the electron beam direction. *c*, Sketch of the centre of *b* with labelled tunnel sizes. The tunnels all have a triple chain of manganese octahedra in common.

magnification of the nodular todorokite material, and they contribute to the streaking of the todorokite diffraction patterns. In some samples, the tunnel structure is disrupted by faults along the chain length. Figure 4*a*, shows a quadruple chain tunnel becoming a triple chain at a fault. The adjacent triple chains are displaced below the fault. Similar disruptions are evident in the HRTEM image of Chukhrov *et al.*¹¹ On a larger scale, narrow fibres are gently warped. A distinct bend in a fibre is seen in Fig. 4*b*. Similar phenomena are found in non-marine todorokite.

In some sample grains, a different phase was seen topotactically intergrown with the todorokite. This other phase gives fringes between 4.5 and 5 Å in width. This spacing may be obscured in X-ray diffraction patterns due to its similarity to a common spacing in todorokite. The identity of this nodule 5 Å phase is uncertain, but structural images obtained from it resemble those we have previously acquired from the double-chain manganese oxides—ramsdellite and groutite. In some cases, todorokite and this 5 Å phase are evident in the same image. In other samples, only the 5 Å phase is seen, and it is associated with apparent relict fibres of todorokite (Fig. 5*a*). In Fig. 5*a*, the large, 100 Å-wide fibres resemble todorokite in the way they are twinned at 120° angles. However, high-resolution imaging shows only amorphous regions and fringes from the 5 Å phase (Fig. 5*b*). It is possible that fibres of todorokite altered to the 5 Å phase. The chemistry of the 5 Å material is

uncertain at present, but preliminary work has indicated that iron is not present. The chemistry will be important in determining the genetic relationship between the todorokite intergrowths and the 5 Å phase.

This study indicates that further use of HRTEM in the study of the cryptocrystalline manganese nodule phases is warranted both to confirm earlier deductions from X-ray work and to find new phases and their detailed structural features. Such structural information should eventually provide clues to long-standing problems regarding the origin and accretion of nodules. For instance, different environments of formation of nodule todorokite may influence their predominant tunnel size. In the nodule todorokite samples of Chukhrov *et al.*^{11–13}, a predominant [100] width is 14.6 Å which corresponds to quintuple chains of manganese octahedra, whereas in our samples the predominant width is of triple chains. Perhaps a layer 10 Å phase (analogous to busserite) occurs in other environments. Two other manganese oxide phases—birnessite and vernadite—have been proposed to occur in nodules. Both these structures are poorly understood and HRTEM should provide insight into their nature. A final problem is the location of transition elements such as Cu, Ni and Co in nodule phases. Use of structural imaging with concurrent microanalysis should firmly establish the identity of the host phases.

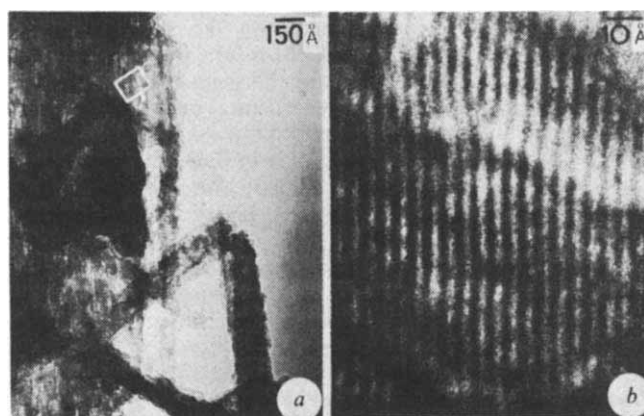


Fig. 5 Images of the possible double-chain phase. *a*, Fibres twinned at 120° are evident. *b*, Close up of box in *a* shows fringes of the possible double-chain phase.

We thank S. Selkirk for aid in drafting Figs 1 and 3*c*, Dr J. E. Post for reviews, and K. Jones and E. Seeling for assistance with the scanning electron microscopy. M.S. thanks Drs H. D. Holland and R. Burns for guidance and support. This research was supported by grants EAR79-26375 and EAR 79-04364 from the Earth Sciences Division of the NSF and the NSF-1D0E Seabed Assessment Program Manganese Nodule Program (grant no. 832550CE). The high-resolution imaging work was performed in the HREM Regional Facility in the Center for Solid State Science at Arizona State University.

Received 14 December 1981; accepted 8 March 1982.

1. Burns, R. G. & Burns, V. M. *Rev. Miner.* **6**, 1–46 (1979).
2. Burns, R. G. & Burns, V. M. in *Marine Manganese Deposits* (ed. Glasby, G. P.) 185–248 (Elsevier, New York, 1977).
3. Straczek, J. A., Horen, A., Ross, M. & Warshaw, C. M. *Am. Miner.* **45**, 1174–1184 (1960).
4. Manheim, F. T. in *Symposium on Marine Chemistry* Vol. 3 (eds Schink, D. R. & Corliss, J. T.) 217–276 (University of Rhode Island, 1965).
5. Buser, W. & Grütter, A. *Schweiz. miner. petrogr. Mitt.* **36**, 49–62 (1956).
6. Giovanoli, R., Feitknecht, W. & Fischer, F. *Helv. chim. Acta* **54**, 1112–1124 (1971).
7. Turner, S. & Buseck, P. R. *Science* **212**, 1024–1027 (1981).
8. Paterson, E. *Am. Miner.* **66**, 424–427 (1981).
9. Turner, S., Siegel, M. & Buseck, P. R. *Trans. Am. geophys. Un.* **62**, 306 (1981).
10. Siegel, M. D. thesis, Harvard Univ. (1981).
11. Chukhrov, F. V., Gorshkov, A. I., Sivtsov, A. V. & Beresovskaya, V. V. *Izv. Akad. Nauk SSSR, Geol.* **12**, 86–95 (1978).
12. Chukhrov, F. V., Gorshkov, A. I., Sivtsov, A. V. & Beresovskaya, V. V. *Nature* **278**, 631–632 (1979).
13. Chukhrov, F. V., Gorshkov, A. I. & Sivtsov, A. V. *Izv. Akad. Nauk SSSR, Geol.* **5**, 88–91 (1981).
14. Siegel, M. D. & Turner, S. *Science* (submitted).
15. Veblen, D. R. & Buseck, P. R. *Am. Miner.* **64**, 687–700 (1979).

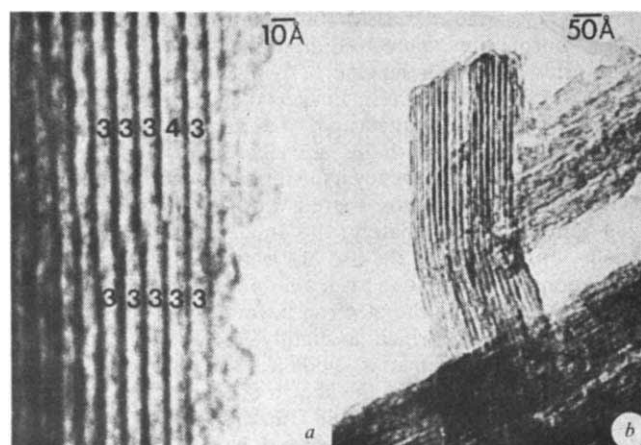


Fig. 4 Disordering phenomena in todorokite. *a*, A disruption along the tunnel length. A quadruple chain tunnel merges to a triple chain. *b*, A bent fibre of todorokite.

Pulsating flow of a plastic fluid

N. Phan-Thien & J. Dudek

Department of Mechanical Engineering, The University of Sydney, New South Wales 2006, Australia

It has been observed^{1,2} that when a polymeric solution flows through a circular pipe under a pressure gradient which varies periodically about a non-zero mean, the mean volumetric flow rate increases over that which is due to the mean pressure gradient alone. The magnitude of this flow enhancement depends on the fluid properties and is a function of the mean pressure gradient and the amplitude and the frequency of its varying component. Theoretical work¹⁻⁶ has indicated that this phenomenon is due to the shear-thinning property of the polymeric fluid. More generally, the flow enhancement arises because of the non-linear relationship between the stress and the strain rate which causes the effective viscosity to decrease. Indeed, we show here that a flow enhancement is possible for a Bingham fluid⁷.

This fluid does not deform below a certain yield shear stress τ_y ; above this yield shear stress it flows like a newtonian fluid (with constant viscosity) and is thus not shear-thinning in the normal sense. However, a power analysis shows that the extra power required to pulsate the flow is greater than that required to generate the extra flow enhancement under a steady pressure drop. Consequently there is no economic reason to pulsate the flow. The set-up of the experimental rig is shown in Fig. 1. A peristaltic pump B was used to circulate the test fluid from a reservoir A which was immersed in a thermal bath. The temperature was regulated to $\pm 1^\circ\text{C}$. Pressure gradient noises in the pump were effectively damped out using two smoothing bottles C in series. The pulsatile pressure gradient was generated by a piston E driven by an offset circular cam D. High frequency components of this pulsatile pressure gradient were damped out using another smoothing bottle F. In this arrangement the fluid flowed under a pressure gradient of the form

$$\partial P/\partial z = -P_0(1 + \varepsilon n(t)) \quad (1)$$

where P_0 is the mean pressure drop per unit length, $n(t)$ is a near perfect sinusoidal function of time and ε is the relative amplitude of the pulsatile pressure gradient. To measure P_0 , a

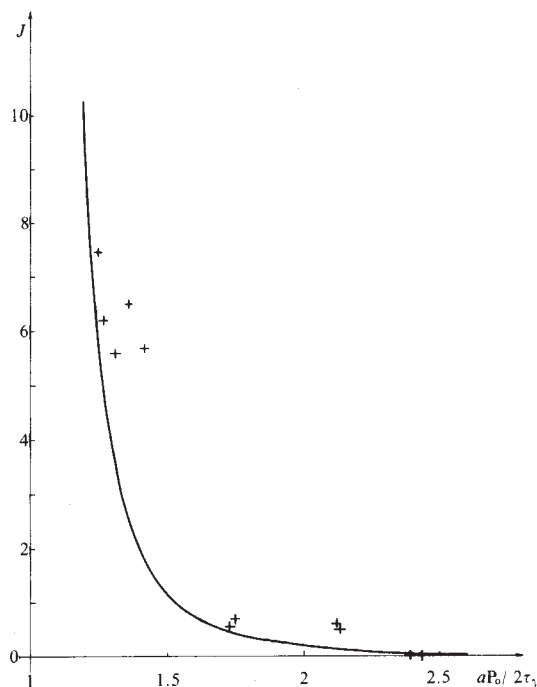


Fig. 2 The base flow enhancement. Solid line is the theoretical curve based on equation (5). +, Data on a 10% bentonite solution taken at a frequency of 0.1 Hz, temperature $(24.5 \pm 0.5)^\circ\text{C}$ and the amplitudes (ε) vary between 0.3 and 0.4.

differential pressure transducer H connected to the test section G at K and L was used. To avoid the attenuation in the amplitude of the pulsatile pressure gradient, we measured ε by mounting the transducer H directly on K and disconnecting the plastic tube J. The efflux was measured by the conventional beaker and stopwatch method.

The test fluid is an aqueous suspension of bentonite at 10% by weight. The bentonite powder (manufactured by Magcobar) conformed to API specification 13A. On a wet screen analysis it leaves a maximum of 4% residue on a US standard sieve no. 200 (nominal opening of $74\ \mu\text{m}$). At low shear rate the bentonite suspension behaves essentially like a Bingham fluid. That is, in a simple shear flow of shear rate $\dot{\gamma}$, the shear stress is given by (note that τ is an odd function of $\dot{\gamma}$)

$$\tau = \eta \dot{\gamma} + \tau_y \text{sign}(\dot{\gamma}) \quad (2)$$

where $\text{sign}(\dot{\gamma}) = 1$ if $\dot{\gamma} > 0$ and $\text{sign}(\dot{\gamma}) = -1$ if $\dot{\gamma} < 0$. Using a commercial vane viscometer (Wykeham Farrance Engineering Ltd) and an Instron Model 3250 Rotary Rheometer with a cone (2°) and plate (2 cm radius), we found that $\tau_y = (6.4 \pm 0.5)$ Pa.

To quantify the effect of the pulsatile pressure gradient, we define a flow enhancement I as

$$I = (\langle Q \rangle - Q_0)/Q_0 \quad (3)$$

where $\langle Q \rangle$ is the average flow rate (the angular brackets denote a time average) and Q_0 is the flow rate due to the mean pressure gradient alone.

At low value of ε ($\varepsilon < 0.4$) and keeping the mean pressure drop and the frequency of the pulsatile pressure gradient fixed, we found that I is proportional to ε^2 , which was already noted by Barnes *et al.*^{1,2} for polymeric liquids. However, for the bentonite solution tested we found I to be independent of the frequency ω of the pulsatile pressure gradient, up to a value of $\omega \sim 1$ Hz. The experimental findings at $\omega = 0.1$ Hz are dis-

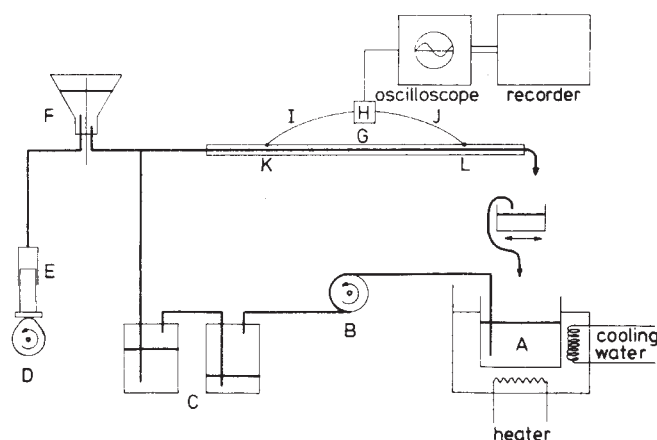


Fig. 1 Schematic arrangement of the pulsating flow experiment.

played in Fig. 2 where we plotted the base flow enhancement $J = I/2\varepsilon^2\langle n^2 \rangle$ against a dimensionless pressure drop $aP_0/2\tau_y$, where $a = 3.2$ mm is the inner radius of the test section. (Note that $2\langle n^2 \rangle = 1$ for a sinusoidal $n(t)$.)

To compare the experimental findings with the theory based on the Bingham fluid model we use equations (1) and (2) in the equations of motion neglecting body forces and the fluid inertia. The latter assumption is reasonable in view of our Reynolds number (based on ω and a) is always <0.01 . It can be easily shown that the flow enhancement is given by

$$I = 2\varepsilon^2\langle n^2 \rangle J + o(\varepsilon^2) \quad (4)$$

where

$$J = (\phi^4 - \frac{4}{3}\phi^3 + \frac{1}{3}), \quad \phi = aP_0/2\tau_y \quad (5)$$

Equations (4) and (5) show that the theoretical flow enhancement is independent of the frequency and that it is proportional to ε^2 for small ε^2 . Furthermore a knowledge of the yield shear stress (and not the viscosity η) and the mean pressure drop per unit length is sufficient to determine the base flow enhancement J . The theoretical base flow enhancement is plotted in Fig. 2. Clearly, in view of a large experimental error which must exist when taking the difference between two quantities of the same order of magnitude, $\langle Q \rangle$ and Q_0 , the agreement between theory and experiment is very satisfactory.

To find out whether pulsating a pipe flow of a Bingham fluid is also economically viable we must also look at the net power requirement. Now the net power required to pulsate the flow is

$$W = - \left\langle Q \frac{\partial P}{\partial z} \right\rangle \\ = \frac{\pi a^4 P_0^2}{8\eta} \left(-\frac{4}{3}\phi^{-1} + \frac{1}{3}\phi^{-4} + \varepsilon^2(1 + \phi^{-4})\langle n^2 \rangle + o(\varepsilon^2) \right)$$

In steady conditions we need a pressure drop per unit length of

$$P' = P_0(1 + 2\varepsilon^2\langle n^2 \rangle(\phi^{-4} - 1) + o(\varepsilon^2))$$

to generate a flow rate $\langle Q \rangle$. Thus the power required in steady-state conditions which yields an identical (to $o(\varepsilon^2)$) flow rate is $W' = P'\langle Q \rangle$. The extra power required to pulsate the flow is

$$\Delta W = W - W' \\ = \varepsilon^2\langle n^2 \rangle(\phi^8 - 4\phi^4 + \frac{8}{3}\phi^3 + \frac{1}{3})(\phi^4 - 1)^{-1} \\ \times (\phi^4 - \frac{4}{3}\phi^3 + \frac{1}{3})^{-1} + o(\varepsilon^2)$$

If terms of $o(\varepsilon^4)$ and higher are neglected, ΔW is seen to be always positive. Consequently there is no economic advantage in pulsating the flow of a Bingham fluid. This conclusion does not agree with that of Barnes *et al.*² who modelled their polymeric fluids by an Oldroyd model.

This research is supported by an ARGC grant (F 79 1559 7R). We thank the Oil Drilling and Exploration Company for donating the bentonite sample.

Received 12 January; accepted 11 March 1982.

1. Barnes, H. A. & Walters, K. *Nature* **216**, 366 (1967).
2. Barnes, H. A., Townsend, P. & Walters, K. *Nature* **224**, 585 (1969); *Rheol. Acta* **10**, 517 (1971).
3. Sundstrom, D. W. & Kaufman, A. *Ind. Eng. Chem. Proc. Des. Dev.* **16**, 320 (1977).
4. Davies, J. M., Bhumiratana, S. & Bird, R. B. *J. non-Newt. Fluid Mech.* **3**, 237 (1978).
5. N. Phan-Thien, *J. non-Newt. Fluid Mech.* **4**, 161 (1978); *Rheol. Acta* **19**, 285 (1980); *J. Rheol.* **25**, 293 (1981).
6. Manero, O. & Walters, K. *Rheol. Acta* **19**, 277 (1980).
7. Bingham, E. C. *Fluidity and Plasticity* (McGraw-Hill, New York, 1922).

Steam explosions of molten iron oxide drops: easier initiation at small pressurizations

Lloyd S. Nelson & Patricia M. Duda

Sandia National Laboratories, Albuquerque, New Mexico 87185, USA

If a light water nuclear reactor core were to overheat, hot molten materials might contact liquid water and cause a steam explosion damaging to the containment structure¹. To predict such an occurrence and the ensuing damage, we in the Sandia National Laboratories have been performing large² and small^{3,4} scale experiments coupled with analytical modelling^{5,6}. As part of this effort, we have been releasing single drops of a core melt simulant, molten iron oxide, into liquid water. Small steam explosions are triggered shortly afterwards by applying a pressure pulse to the water. The threshold peak pulse level above which an explosion always occurs was studied at ambient pressures between 0.083 and 1.12 MPa. Unexpectedly, the threshold decreased to a minimum in the range 0.2–0.8 MPa and then increased again. The observation that gentle pre-pressurization meant easier triggering should improve the predictive capability of our models and suggest conditions for reducing the hazards of these explosions.

A steam explosion can occur when a hot, molten material contacts liquid water. The explosion, which sometimes generates destructive shock waves, is caused by the fragmentation of the melt and rapid transfer of heat to the water with very rapid generation of steam. Explosions such as these pose serious hazards whenever large quantities of molten materials are handled, for example, in the metallurgical and papermaking industries. (Cronenberg and Benz⁷ recently reviewed steam explosions with particular emphasis on nuclear reactor safety.)

Melts that would result from a nuclear reactor core meltdown, called coriums⁷, would be composed of various mixtures of UO₂ fuel, zircaloy cladding, and stainless steel structures, oxidized to some extent by contact with steam at temperatures near 2,000 K. We have been studying iron oxide because it can be a major component of some of the coriums⁸; it also simulates some of their physical and chemical properties. Single drops are used because it enables excellent control of the many parameters involved in the interactions.

Individual 2.9-mm diameter pendant drops of iron oxide were heated to 2,230 K with 100–200 W of focused continuous wave CO₂ laser radiation^{4,9}. The drops, supported on the bottom end of a vertical iridium wire, were positioned a few millimetres above the surface of liquid water held in a polymethylmethacrylate tank 152 mm square and 300 mm deep. The tank contained a water heater, stirrer and thermometer. At various distances from the bottom of the tank, a pressure pulse was generated in the water by the capacitor discharge explosion of a submerged wire. The magnitude of the pulse which triggered the steam explosion was varied by raising or lowering the exploding wire. The interactions were analysed using high speed cameras, a lithium niobate pressure transducer^{10,11} hung in the water, and by examination of the iron oxide residues after the interactions.

The entire apparatus was enclosed in a steel pressure chamber which had a zinc selenide window¹², a transparent fused silica window, and a thick glass window for passage of the laser radiation, pyrometer viewing and photography, respectively. The oxygen partial pressure within the chamber was held at ≈ 0.02 MPa to maintain the composition of the drop at FeO_{1.2} (ref. 13). The total chamber pressure was varied from local atmospheric (0.083 MPa) to 1.12 MPa by pressurizing with Ar or N₂. After shaking the molten drop from the wire, the interac-

tion was initiated by exploding the wire ≈ 125 ms after the drop entered the water.

The steam explosion produces a steam bubble which alternately grows and collapses for several complete cycles. Each bubble collapse produces a pressure pulse in the water. These pulses, along with the triggering pulse, cause progressively finer fragmentation of the melt, presumably due to collapse of the film boiling layer that surrounds each melt particle at the beginning of each cycle. Bubble growth is caused by the rapid transfer of melt enthalpy to the water with each breakup of the melt. The final debris consists of finely divided iron oxide with average particle diameters in the tens of micrometres range. Further description of the experiments is given in refs 4 and 14.

Recently, we have studied the effects of varying the basic parameters involved in the interactions¹⁴, including ambient pressure. The effects of increased ambient pressure on vapour explosions have been considered previously, both analytically^{6,15-18} and experimentally (several experimental reports are cited in ref. 6). It seems established that for given triggering conditions, it should become progressively more difficult and eventually impossible to initiate an explosion as ambient pressure is increased. However, it has been neither predicted nor observed that initially it becomes easier before it becomes harder to trigger an explosion as this pressure increases, at least for single drops of a melt such as ours.

The easier triggering was first observed as a sharp drop in the threshold triggering peak pressure, Δp , required to initiate an explosion as ambient pressure, P_∞ , was increased. This is shown in Fig. 1, which indicates the occurrence or absence of an explosion as these two parameters were varied; water temperature was held constant at 297 K. Note the abrupt dip in threshold trigger pressure from $\Delta p \approx 0.4$ to <0.2 MPa as P_∞ is increased from 0.083 to 0.2 MPa, and the broad minimum in Δp which extends to $P_\infty \approx 0.8$ MPa.

As chamber pressure is increased at constant water temperature, the water subcooling also increases. (The subcooling of a liquid is the difference between its boiling temperature at a given ambient pressure and its actual temperature. Subcooling is normally considered to be one of the parameters which governs heat transfer in film boiling situations of the sort that prevail in our experiments.) Because at a given ambient pressure

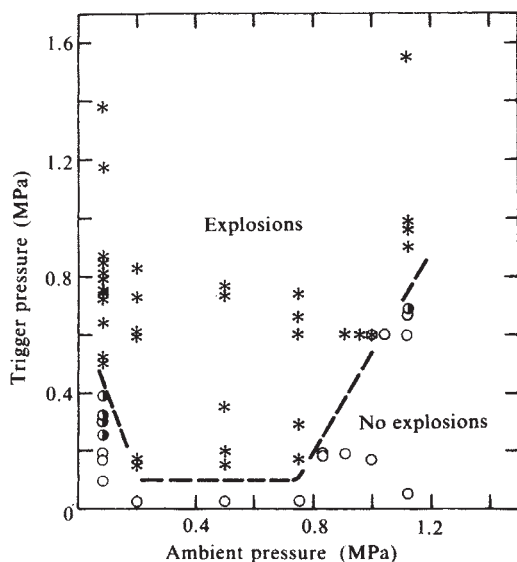


Fig. 1 Plot of peak triggering pulse pressure against ambient chamber pressure for the interaction of a single 2.9 mm drop of molten iron oxide at 2,230 K released into water. Temperature of the water was 298 K for the points at an ambient pressure of 0.083 MPa; for all other experiments, it was 296 ± 1 K. *, Vigorous explosion with fine fragmentation. ●, Sluggish interaction, often with an explosion ≈ 0.1 s after the triggering pulse. ○, No explosion.

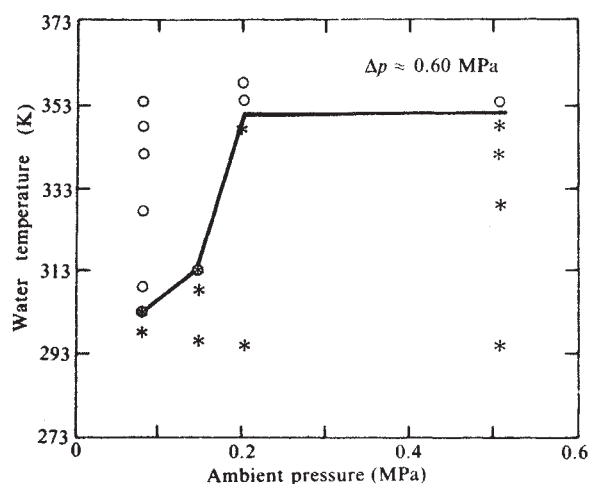


Fig. 2 Effect of water temperature on the interaction between single drops of molten iron oxide and liquid water as ambient pressure, P_∞ , varies. A constant peak trigger pressure, Δp , of ≈ 0.60 MPa is used in all experiments. *, Vigorous explosion with fine fragmentation. ○, No explosion. ●, Duplicate experiments: one drop exploded, the other did not.

steam explosions normally become easier to initiate as water subcooling increases⁶, the question arises as to whether the increased ease of initiation as ambient pressure increases is due primarily to water subcooling.

We therefore performed a series of experiments whereby ambient pressures were increased but water temperatures were also increased appropriately to maintain constant water subcooling. We again found that at an ambient pressure of 0.2 MPa the trigger level required dipped to <0.3 MPa; however, it increased more rapidly than in Fig. 1 to ≈ 0.7 MPa at $P_\infty = 0.5$ MPa. The increased ease of triggering a steam explosion at increased ambient pressures, therefore, seems to be due only partially to water subcooling effects.

We then investigated the easier triggering as pressure increased by varying water temperature and ambient pressure in a different way, keeping the triggering level fixed at $\Delta p \approx 0.6$ MPa. Because it becomes more difficult to initiate steam explosions as the water becomes hotter at constant ambient pressure⁶, the threshold temperature above which explosions no longer can be initiated for a fixed triggering level becomes another measure of the ease of initiating a given melt-water interaction.

The measured threshold temperatures shown in Fig. 2 were 300 and 310 K for $P_\infty \approx 0.083$ and 0.14 MPa, but rose sharply to 350 K for $P_\infty \approx 0.20$ and 0.50 MPa. This sharp rise again indicates that triggering the steam explosion of a single drop of molten iron oxide becomes easier at ambient pressures a few times greater than atmospheric.

Although the effect is as yet unexplained we feel that the easier initiation as ambient pressure increases may have an important role in the triggering and propagation of a large scale steam explosion through a coarsely premixed dispersion of melt in water. For example, if a quasistatic localized pressurization should occur in the early stages of the explosion (water temperature would probably remain essentially constant), this would increase the sensitivity of the pressurized portion of the array to subsequent pressure disturbances generated in the system. This may aid the transmittance of the explosive interaction throughout the system.

This work was supported by the US Nuclear Regulatory Commission.

Received 27 October 1981; accepted 19 February 1982.

1. Cronenberg, A. W. & Benz, R. *Adv. Nucl. Sci. Technol.* **12**, 247-334 (1980).
2. Mitchell, D. E., Corradini, M. L. & Tarbell, W. W. *Intermediate Scale Steam Explosion Phenomena: Experiments and Analysis* (Sandia National Laboratories, Albuquerque, SAND81-0124, 1981).

3. Nelson, L. S. *High Temp. Sci.* **13**, 235-256 (1980).
4. Nelson L. S. & Duda, P. M. *Steam Explosion Experiments With Single Drops of Iron Oxide Melted With a CO₂ Laser* (Sandia National Laboratories, Albuquerque, SAND81-1346, 1981).
5. Corradini, M. L., Swenson, D. V., Woodfin, R. L. & Voelker, L. E. *Nucl. Engng Design* **66**, 287-298 (1981).
6. Corradini, M. L. *Nucl. Sci. Engng* **78**, 154-170 (1981).
7. Peehs, M. *Investigations of Molten 'Corium' Phases* (IAEA, Vienna, IAEA-SM-190/10, 1975).
8. Nazare, S., Ondracek, G. & Schulz, B. *Nucl. Technol.* **32**, 239-246 (1977).
9. Nelson, L. S., Richardson, N. L., Kiel, K. & Skaggs, S. R. *High Temp. Sci.* **5**, 138-154 (1973).
10. Graham, R. A. *Lithium Niobate Stress Transducers*, 19-23 (Sandia National Laboratories, Albuquerque, SAND75-0426, 1975).
11. Graham, R. A. & Jacobson, R. D. *Appl. Phys. Lett.* **23**, 584-586 (1973).
12. Nelson, L. S. & Duda, P. M. *Rev. Scient. Instrum.* **53**, 259-260 (1982).
13. Darken, L. S. & Gurry, R. W. *J. Am. chem. Soc.* **68**, 798-816 (1946).
14. Berman, M. *LWR Safety Research Program Q. Rep.* (Sandia National Laboratories, Albuquerque, SAND80-1304, 1980; SAND81-1216, 1981).
15. Buchanan, D. J. *J. Phys. D.: Appl. Phys.* **7**, 1441-1457 (1974).
16. Buchanan, D. J. & Dullforce, T. A. *Nature* **245**, 32-34 (1973).
17. Peckover, R. S., Buchanan, D. J. & Ashby, D. E. T. F. *Nature* **245**, 307-308 (1973).
18. Henry, R. E. *Test Plan—Large Scale Molten Salt-Water Vapour Explosion Studies to be Conducted at Ispra, Italy* (Argonne National Laboratories, Argonne, ANL-79-20, 1978).

Optical spectroscopy in a shocked liquid

G. E. Duvall*, K. M. Ogilvie*, R. Wilson*,
P. M. Bellamy* & P. S. P. Wei†

*Department of Physics, Washington State University, Pullman, Washington 99164-2814, USA

†Boeing Aerospace Company, Seattle, Washington 98124, USA

Scientists in the USSR have reported chemical effects of shock waves in condensed materials which they cannot attribute to the effects of pressure and temperature. They postulate the existence of "catastrophic" chemical effects in the region of near-discontinuous change called the "shock front"^{1,2}. Direct testing of this postulate requires time-resolved observations of chemical processes as the sample is traversed by the shock, and to this end we report here the development of methods for the measurement of temporal changes in the electronic spectra of condensed materials in such conditions. The time resolution is about 3×10^{-8} s.

Measurements in liquid CS₂ show unexpected increases in reflectivity and absorption as well as large shifts of absorption band edges towards the red at shock pressures of <12 kbar. The significance of these observations is not yet understood, but it is clear that electronic, and therefore chemical, changes are taking place in transient conditions not normally accessible to chemical experimentation.

In 1920, Parsons reported unsuccessful attempts to produce artificial diamonds by compression of carbonaceous materials with shock waves³. Forty-one years later the feasibility of the conversion was established⁴. Work before 1961 had shown that profound changes could be produced in metals by strong shock waves; later experiments, carried out principally in the USSR, have demonstrated the feasibility of driving a variety of chemical reactions by compression with moderate to strong shock waves^{1,2}. While these experiments, called 'recovery experiments' because analysis takes place only on samples collected after being shocked, were in progress, various types of dynamic shock experiments were developed and used for studying mechanical, thermal, electrical and optical properties of condensed matter⁵. These latter experiments are most successful if plane one-dimensional shock waves are used to compress the samples. Such shock waves can be produced in the laboratory by planar impact of a projectile on the sample.

The oldest known chemical reaction initiated by shock waves is detonation in high explosives. Several years ago, moved by the difficulty of making dynamic measurements in detonation, Sheffield undertook dynamic measurements on a non-explosive

material, liquid CS₂, to establish its decomposition, to provide evidence that dynamic measurements in such a system are possible and to measure kinetic parameters of the process^{6,7}. The success of his measurements led to a desire for measurements more directly related to chemical processes and thus to the development of spectroscopic techniques reported here.

Optical spectroscopy in shock experiments has been attempted before but with qualified success⁸⁻¹¹. In our experiments, a liquid sample is contained in a penny-shaped cell, bounded front and back by transparent sapphire disks and on the sides by a cylindrical brass shell. The shock is created by impact of a flat-faced projectile on the plane face of one of the sapphire disks. Just before impact a xenon flash lamp is pulsed, producing a spectrally broad light which traverses the sample assembly, passes through a grating spectrograph and is recorded on 35 mm film by means of a rotating mirror streak camera. The record on the film has wavelength varying across the film strip, time varying along the strip and density varying with reflectivity or transmissivity of the sample. For reflection experiments the light source is external to the target and projectile assemblies. Light is brought into the cell through the sapphire disk opposite the impact side and in a direction very close to normal. It comes out near the normal and is reflected through the entrance slit of the spectrograph. For transmission experiments the flash lamp is mounted in the projectile. Contacts on the face of the projectile connect with high voltage leads in the target just before impact. The power source is triggered soon afterwards and light transmitted through the cell is reflected

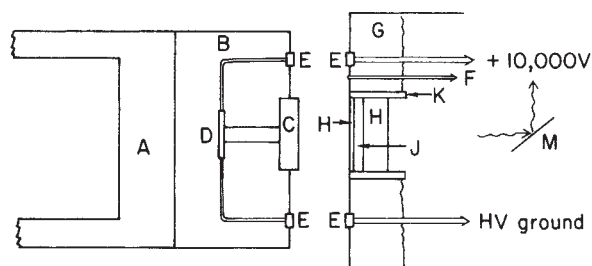


Fig. 1 Schematic representation of light transmission experiment. A, aluminium body of projectile; B, PMMA/epoxy projectile head; C, sapphire impactor; D, Xe flash lamp; E, high-voltage connectors; F, trigger pin for oscilloscopes; G, epoxy target body; H, sapphire plates of test cell; J, cavity containing liquid CS₂; K, brass container for test cell; M, mirror which reflects light into objective mirror of spectrograph.

into the spectrograph as before. The subject material has been CS₂ at room temperature and an initial pressure of 1 atmosphere. Thicknesses of the CS₂ layer vary from ~150 μm to <1 μm, and final pressure in the sample is reached by reverberation between the sapphire plates. Reflection records obtained thus far are in the UV, approximately centred on the transmission region at 2,700 Å between two absorption bands¹². The records show transmission as expected before the shock reaches the sample and momentary extinction as the shock enters the sample, followed by reflection all across the film as pressure in the sample rings up to its final value. Reflection coefficients are estimated to be >0.2 at 80 kbar and >0.6 at 120 kbar. At a final pressure of 80 kbar, reflection remains constant after the first three or four reverberations. At 120 kbar, it reaches a peak and then decays with a time constant of about 280 ns. Transmission experiments show a shift of absorption edges towards the red at an initial shock pressure of ~5 kbar by amounts which cannot be explained simply as thermally enhanced excitation of normal vibrational states. At a pressure of <20 kbar, extinction through the range 2,500-4,000 Å is complete. Transmission in the visible is almost completely extinguished at 120 kbar in this reverberation geometry.

A schematic diagram of the transmission experiment is shown in Fig. 1. The projectile is 10.16 cm in diameter and 20.32 cm long, driven by high-pressure gas¹³. The reflection experiment differs from Fig. 1 in that the flash lamp is replaced by a cavity behind the impactor and the left face of the sapphire impactor is coated with a light absorber to minimize two-way transmission through the cell. When two-way transmission is desired the left face can be made a mirror. With the reflector in this position, it is unshocked during the significant part of the experiment.

A transmission record is shown in Fig. 2. The region 3,580–2,900 Å is an absorption band; there is another below 2,680 Å, but its edge is obscured by decreasing sensitivity of film and grating. At impact the reflection from air–sapphire interfaces between target and impactor is reduced, causing the barely perceptible increase in density across the film at *a*. At *b*, the shock initiated at impact has traversed the 2 mm face of the cell and entered the CS₂. At *c*, the shock in CS₂ has traversed the cell for the first time and reflected from the inside face of the back sapphire plate. The strength of the first shock is ~5 kbar. On reflection this increases to 12 kbar, which is apparently enough to cause complete extinction of the transmitted light in the lower transmitting region. Temperatures for these first two shocks were calculated to be 398K and 482K respectively; the apparent change in the absorption edge which is initially below 2,670 Å is >0.23 eV; for the upper edge at 3,580 Å the change is 0.11 eV for the first shock. The shift due to the second shock in the upper edge is ~0.12 eV. Transmission in the visible (not shown) was little affected at the final pressure of ~55 kbar.

The record from a reflection experiment is shown in Fig. 3. In this experiment the left-hand face of the sapphire impactor, C in Fig. 1, is silvered, there is no flash lamp, and the record to the left of *b* in Fig. 3 represents two-way transmission through the CS₂ sample. The first shock, with amplitude of 8 kbar, enters the CS₂ at point *b* and completes its traverse of the CS₂ at *c*. The extinction of light during this interval is similar to that recorded in Fig. 2, but pressure of the first shock is larger, ~8 kbar, and calculated temperature is 427K. Reflection of the shock at the second CS₂ sapphire interface at *c* initiates

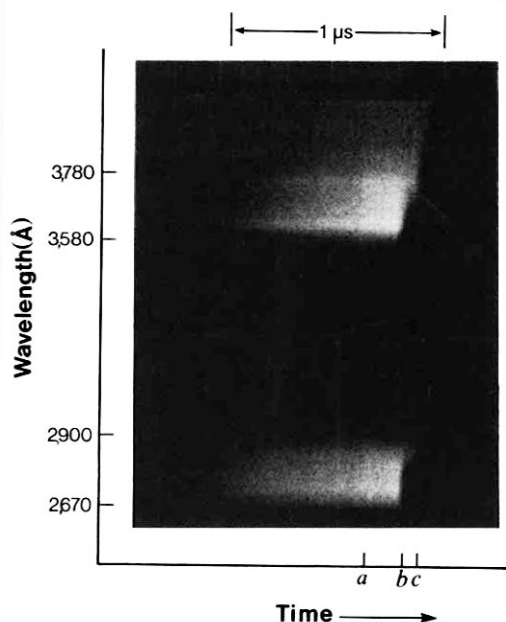


Fig. 2 Time-resolved spectrogram of light transmission through the CS₂ filled cell shown in Fig. 1. Light regions represent transmission. *a*, impact occurs; *b*, shock arrival at sapphire–CS₂ interface; *c*, second shock traversal of CS₂ cell (first reverberation). CS₂ sample is 145 μm thick and ~3 cm in diameter. Time resolution is ~30 ns; spectral resolution is ~30 Å. Final pressure is 55 kbar. Shot number 81-010. ×3.2.

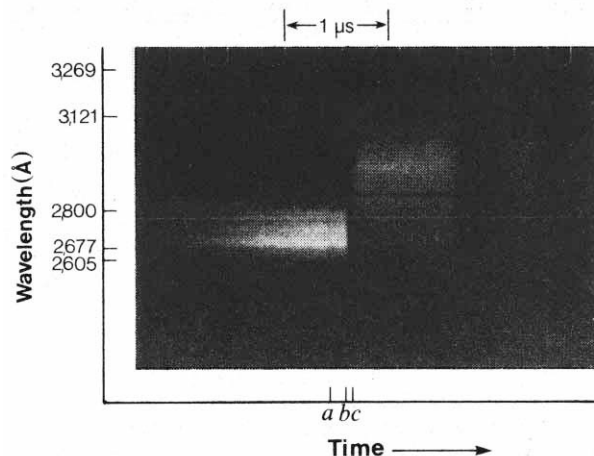


Fig. 3 Time-resolved spectrogram of two-way transmission through and reflection from a CS₂-filled cell. *a*, Impact occurs; *b*, shock reaches front interface of CS₂–sapphire; *c*, shock completes first traverse of CS₂ cell. Cell thickness is 69 μm. Time and spectral resolutions as in Fig. 2. Final pressure is 80 kbar. Shot number 80-010. ×1.6.

the optical reflection process. The strength of the reflection in the next few reverberations builds up to a value which remains constant until the shock reaches the rear face of the second sapphire. This is corroborated by a photomultiplier record (not shown) of the total light intensity plotted against time, integrated over the wavelengths indicated in Fig. 3. Final temperature at 80 kbar is ~870K.

Measurements of resistivity of the CS₂ in these shock conditions give values of the order of 10⁺⁴ Ωm, which is too large to explain either the observed extinction or reflection. Transmission spectra in the visible show that transmission for wavelengths ≥4,000 Å persists at a final (ring-up) pressure of 55 kbar. This confirms qualitative experiments reported elsewhere¹⁴.

Final pressure in these experiments is reached by a series of reverberations, so they are not directly comparable to single shock experiments to the same pressure. Final temperature is less in the reverberation cell than in a single shock.

These experiments should allow processes of molecular interaction which have heretofore been inferred from indirect evidence to be observed directly and open new avenues for studying the dynamics of chemical processes in highly compressed matter. Further details can be found in ref. 15.

We acknowledge the continuing support of this research by the United States Office of Naval Research under contract no. N00014-77-C-0232 and the encouragement provided by Dr Richard S. Miller.

Received 15 December 1981; accepted 29 February 1982.

- Graham, R. A. in *Proc. int. Conf. on Metallic Effects of High Strain Rate Deformation and Fabrication*, Albuquerque (1980).
- Dremin, A. N. & Breusov, O. N. *Priroda* No. 12, 10–17 (1971). (Transl. SAND-80-6003, 1980).
- Parsons, C. A. *Phil. Trans. R. Soc. A* **220**, 67–108 (1920).
- DeCarli, P. S. & Jamieson, J. C. *Science* **133**, 1821–22 (1961).
- Davison, L. & Graham, R. *Phys. Rep.* **55**, 256–379 (1979).
- Sheffield, S. A. thesis, Washington State Univ. (1978).
- Sheffield, S. A. & Duvall, G. E. *Proc. Symp. on High Dynamic Pressures* (CNES, Paris, 1978).
- Duvall, G. E. & Moore, D. B. *Poulter Laboratories Technical Rep. No. 017–59* (Stanford Research Institute, California, 1959).
- David, H. G. & Ewald, A. H. *Austr. J. appl. Sci.* **11**, 317–320 (1960).
- Gaffney, E. S. & Ahrens, T. J. *J. geophys. Res.* **78**, 5942–5953 (1973).
- Goto, T., Ahrens, T. J. & Rossman, G. R. *Phys. Chem. Miner.* **4**, 253–263 (1979).
- Herzberg, G. *Electronic Spectra and Electronic Structure of Polyatomic Molecules*, 504 (Van Nostrand, New York, 1967).
- Fowles, G. R. *et al. Rev. Scient. Instrum.* **41**, 984–996 (1970).
- Yakusheva, O. B., Yakushev, V. V. & Dremin, A. N. *Russ. J. phys. Chem.* **51**, 973–975 (1977).
- Ogilvie, K. thesis, Washington State Univ. (1982).

Lipids in Harmattan aerosols of Nigeria

R. E. Cox*, M. A. Mazurek† & B. R. T. Simoneit‡

* Chemistry Department, University of Jos, PMB 2084, Jos, Nigeria

† Department of Earth and Space Sciences and

‡ Institute of Geophysics and Planetary Physics, University of California, Los Angeles, California 90024, USA

Terrestrial lipid detritus has been found in some pelagic sediments at great distances from land (for example, in the North Atlantic) and aeolian transport has been invoked as one input mechanism to explain this observation¹. This is supported by the presence of the same lipids in oceanic aerosols, presumably derived from major wind systems originating on land^{2,3}. Thus, it is vital to examine whether this lipid signature is observed on the continent in the ambient aerosols towards their source regions (such as the Harmattan). We have analysed solvent-extractable lipids in Harmattan aerosols collected in Nigeria and report here that the distributions of hydrocarbons indicate both a major contribution of plant waxes and lesser components of a complex composition, believed to originate from algal and other microbiological detritus in desiccated lacustrine muds or from eroded, recycled sedimentary organic matter.

The origins, nature and fate of Saharan dusts have been discussed elsewhere⁴. Organic material in dust samples collected off the coast of north-west and south-west Africa has been analysed^{2,3} but most of the samples are believed to have a different source from the Harmattan dusts. The Harmattan aerosols are believed to arise in the Faya Largeau area of the Chad Basin⁵ and are carried by the prevailing winds over Nigeria and other parts of West Africa to the Atlantic Ocean (Fig. 1). Analysis of organic matter in dusts collected off the west African coast has indicated a high concentration of organic detritus³. For a better understanding of the nature and origin of the organic matter, it was desirable to sample the dust aerosols closer to their presumed origin. High-volume air samplers with quartz fibre filters capable of collecting particulates down to $\sim 1 \mu\text{m}$ diameter were used⁶. Samples were collected in varying conditions of visibility at several locations in Nigeria (Fig. 1) over two Harmattan seasons. A major problem in organic analysis of aerosols is distinguishing between natural and anthropogenic materials. Anthropogenic components could be present in the local atmosphere and would originate from emissions by vehicle exhausts, kerosene lamps and stoves.

The locations, sampling dates and analytical data for the samples that we analysed are presented in Table 1. Total solvent-extractable organic matter (lipids)⁶ was $600\text{--}6,000 \text{ ng m}^{-3}$ of air, of which $100\text{--}2,000 \text{ ng m}^{-3}$ consisted of hydrocarbons and the total loadings of aerosol were typically $300\text{--}900 \mu\text{g m}^{-3}$ (Table 1). Data are shown for samples collected on typically dusty days at a site in Jos (samples 1–4). This site was chosen to minimize both the effects of local soil dust (elevation 10 m above ground) and anthropogenic interferences (outside the town and upwind of the prevailing northeasterly winds). Further samples were taken in Jos (number 6) and downwind from Jos (number 12). The samples from Sokoto and Kano were taken upwind from the cities and at Maiduguri in the south-west sector (wind from north-east). The samples from Ibadan were taken within the city.

Examples of typical gas chromatographic-mass spectrometric (GC-MS) total ion current traces, equivalent to gas chromatograms, for total hydrocarbon fractions are shown in Fig. 2. An

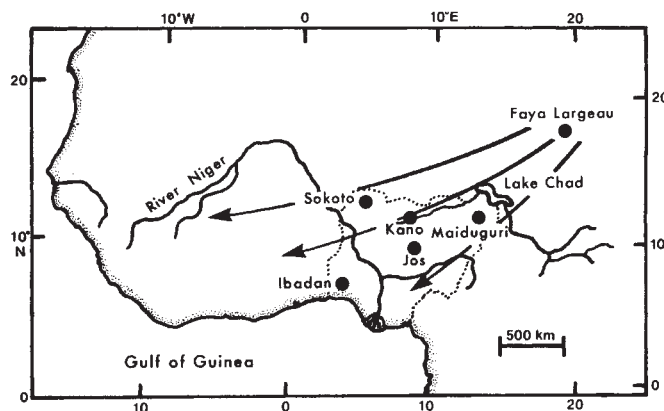


Fig. 1 Location map of the sampling sites and Harmattan wind in West Africa.

outstanding feature of sample 1, upwind from Jos is the value of the carbon preference index (CPI)^{6,7} of 7.8 for the higher *n*-alkanes exhibiting the typical plant wax predominance of odd carbon number alkanes and maximum at *n*-C₂₉ (ref. 8). For comparison, the typical CPI range for various plant waxes is ≥ 5 (refs 8, 9). Also prominent is a complex unresolved mixture (hump) of hydrocarbons centred at the relative retention time of carbon number approximately 19.5. These narrow humps are typically present in degraded detritus of algae and other microorganisms^{10–15} and could have such an origin here. An alternative source may be a recycled component of hydrocarbons (both hump and alkanes) eroded from sedimentary outcrops in the regions where the aerosols are generated. Such recycled material has, for example, been reported in recent sediments in a Swiss lake¹⁶. An unambiguous origin for the narrow hump in these aerosols can therefore not be assigned, based solely on hydrocarbon data. The isoprenoid alkanes 2,6,10,14-tetramethylpentadecane (pristane) and 2,6,10,14-tetramethylhexadecane (phytane) are also present as minor components. The ratio of hump to *n*-alkanes is low (1.4), confirming a relatively low contribution of anthropogenic organic matter, which would be expected to contribute a high proportion of aromatic and unresolvable materials. Based on GC-MS analyses, the amount of aromatic compounds in this sample is very low. This type of compound distribution is observed for samples 1–4 (Table 1).

Data are also shown (Table 1, samples 5 and 6) for aerosol samples taken on moderately clear days in town and in south-west Jos (at the town limit). These show values for alkane CPI intermediate and ratio of hump to normals equal to those of the samples taken upwind of Jos compared with that taken at Ibadan. This organic matter is probably a mixture of natural and petroleum-derived materials.

In contrast, an aerosol sample collected on a day of moderate visibility in the large city of Ibadan (sample 10) shows a high contribution of petroleum residues. This is indicated by a CPI value of 1.7 and the appearance of a new hump of hydrocarbons centred at the relative retention time of carbon number approximately 28.5 (Table 1, Fig. 2b). The ratio of hump to *n*-alkanes is high (4.3). CPI values typical for petroleum are ~ 1.0 and ratios of unresolved (hump) to resolved components are high^{6,17}. The petroleum origin has been further confirmed by GC-MS analyses, which demonstrated the presence of the characteristic molecular markers⁶. Gas chromatograms of hydrocarbon fractions of vehicle exhaust emissions show unresolved humps centred at the relative retention time of the carbon number range C₂₄–C₂₈ (refs 6, 9, and B.R.T.S. unpublished data). This type of contamination is also observed for samples 5 and 6 in Jos and Maiduguri (number 9, Table 1).

Normal fatty acids are present in all samples ($80\text{--}1,000 \text{ ng m}^{-3}$) and exhibit a high even-to-odd carbon number

Present addresses: * School of Natural Resources, University of the South Pacific, PO Box 1168, Suva, Fiji; ‡ School of Oceanography, Oregon State University, Corvallis, Oregon 97331, USA.

Table 1 Location, dates and analytical data for Harmattan aerosols of Nigeria

Sample No.	Location	Lat. Long.	Altitude (m)	Date collected	Visibility (~km)	Volume sampled (m ³)	Total extract (ng m ⁻³)	Total HC (ng m ⁻³)	Hump ÷ normal	Hump max. at C no.	n-Alkanes CPI* (C ₂₅ -C ₃₄)	Max. at C no.†
1	Jos 3 (upwind)	9°56' N 8°54' E	1,200	9 March 1979	2	720	2,200	1,400	1.4	19.5	7.8	29
2	Jos 5 (upwind)	9°56' N 8°54' E	1,200	12 March 1979	8	780	660	240	1.6	24	5.9	29
3	Jos 8 (upwind)	9°56' N 8°54' E	1,200	7 December 1979	2.5	720	1,700	110	2.0	20.5	3.7	29
4	Jos 11 (upwind)	9°56' N 8°54' E	1,200	4 February 1980	1	930	2,400	150	1.4	20	5.7	29
5	Jos 6 (in town)	9°56' N 8°54' E	1,200	22 March 1979	8	670	6,600	2,600	7.7	28.5	3.3	29
6	Jos 12 (downwind)	9°56' N 8°54' E	1,200	14 February 1980	4	1,740	3,300	820	4.0	28.5	2.4	29
7	Kano	12°00' N 8°30' E	520	5 February 1980	1	~600	~6,000	~1,700	2.6	28.5	3.0	29
8	Sokoto	13°03' N 5°14' E	290	22 February 1980	8	1,610	1,200	210	4.0	28.5	4.5	29, 31
9	Maiduguri	11°49' N 13°09' E	290	30 January 1980	8	2,650	5,700	1,400	3.9	29	2.9	29, 31
10	Ibadan	7°23' N 3°53' E	220	25 January 1980	8	3,640	6,000	840	4.6	28.5	1.7	29

* Carbon preference index, summed from C₂₅ to C₃₄, odd divided by even.

† Dominant homologue is underlined.

predominance (CPI=4–11) with maxima at C₁₆, C₂₆ and C₂₈, typical of plant waxes and other biogenic sources¹⁸. Minor amounts of anteiso-fatty acids (C₁₃–C₁₇) and traces of iso-fatty acids are part of the fraction. They are commonly found in microbial lipids¹⁹, further supporting a possible recent biological origin for the hump in the hydrocarbon fraction. Normal fatty alcohols are also present in all samples (230–3,000 ng m⁻³) and exhibit a high even-to-odd carbon number predominance (CPI=4–11) and maxima at C₂₈ and/or C₃₀, again typical of plant waxes and other minor biogenic sources¹⁸. These compounds are not found in petroleum with these distributions and concentrations.

We believe that the samples of Harmattan aerosols collected upwind of Jos, Kano and Sokoto contain largely natural organic matter carried with the dusts or adsorbed on the particulates. The lipid material is composed of primarily plant waxes and secondary compounds derived from degraded microbiological detritus (probably mainly of algae) and/or possibly from eroded sedimentary organic matter. In a study of Harmattan dust

deposition in Nigeria, remobilization of existing dust has been invoked to explain a lack of correlation of dust deposition which visibility and solar radiation in the latter half of the season²⁰. Our samples were collected mainly during the latter half of the dust seasons. The lipid material in our samples may arise predominantly in the Sahara like the dust itself. An alternative and more likely explanation is that the lipids originate predominantly in more fertile areas of Chad and Nigeria downwind of the source. Plant wax lipids could be released by a sand blasting effect of the dust on plant surfaces, while the minor microbial detritus could arise by the remobilization of dust from dried-up wadis and lakes, with a possible eroded sedimentary component.

Further such analyses will permit the tracking of aerosol parcels by their lipid signature back to their respective source regions, and will complement the corresponding mineralogical characteristics.

We thank Dr. C. Ekweozor for assistance with sampling in Ibadan, Dr Leon C. Jacobson for assistance with sampling in Maiduguri and Mr E. Ruth for GC-MS data. We also thank the NSF, Atmospheric Research Section (grant no. ATM 79-08645) for financial support.

Received 9 July 1981; accepted 16 February 1982.

- Farrington, J. W., Frew, N. M., Gschwend, P. M. & Tripp, B. W. *Estuar. Coast. mar. Sci.* **5**, 793 (1977).
- Lepple, F. K. & Brine, C. J. *J. geophys. Res.* **81**, 1141–1146 (1976).
- Simoneit, B. R. T. *Mar. Chem.* **5**, 443–464 (1977). Simoneit, B. R. T., Chester, R. & Eglinton, G. *Nature* **267**, 682–685 (1977).
- Morales, C. (ed.) *Saharan Dust, Scope Rep. 14* (Wiley, New York, 1971).
- Hamilton, R. A. & Archbold, J. W. *Q. Jl R. met. Soc.* **71**, 231–265 (1945).
- Simoneit, B. R. T., Mazurek, M. A. & Cahill, T. A. *J. Am. Pollut. Contr. Assoc.* **30**, 387–390 (1980).
- Cooper, J. E. & Bray, E. E. *Geochim. cosmochim. Acta* **27**, 1113–1127 (1963).
- Eglinton, G. & Hamilton, R. J. in *Chemical Plant Taxonomy* (ed. Swain, T.) 187–217 (Academic, New York, 1963).
- Simoneit, B. R. T. & Mazurek, M. A. *Atmos. Envir.* (in the press).
- Jones, J. G. *J. gen. Microbiol.* **59**, 145–152 (1969).
- Johnson, R. W. & Calder, J. A. *Geochim. cosmochim. Acta* **37**, 1943–1955 (1973).
- Powell, T. G., Cook, P. J. & McKirdy, D. M. *Am. Ass. petrol. Geol. Bull.* **59**, 618–632 (1975).
- Hatcher, P. G., Simoneit, B. R. T. & Gerchakov, S. M. in *Advances in Organic Geochemistry 1975* (eds Campos, R. & Goni, J.) 469–484 (ENADIMSA, Madrid, 1977).
- Cranwell, P. A. *Geochim. cosmochim. Acta* **42**, 1532–1532 (1978).
- Simoneit, B. R. T., Halpern, H. I. & Didyk, B. M. in *Biogeochemistry of Ancient and Modern Environments* (eds Trudinger, P. A., Walter, M. R. & Ralph, B. J.) 201–210 (Springer, New York, 1980).
- Giger, W., Schaffner, C. & Wakeham, S. G. *Geochim. cosmochim. Acta* **44**, 119–129 (1980).
- Tissot, B. P. & Welte, D. H. *Petroleum Formation and Occurrence* (Springer, Berlin, 1978).
- Simoneit, B. R. T. in *Chemical Oceanography*, Vol. 7, 2nd edn (eds Riley, J. P. & Chester, R.) (Academic, London, 1978).
- Boon, J. J., de Lange, F., Schuyf, P. J. W., de Leeuw, J. W. & Schenck, P. A. in *Advances in Organic Geochemistry 1975* (eds Campos, R. & Goni, J.) 255–272 (ENADIMSA, Madrid, 1977).
- McTainsh, G. *Nature* **286**, 587–588 (1980).

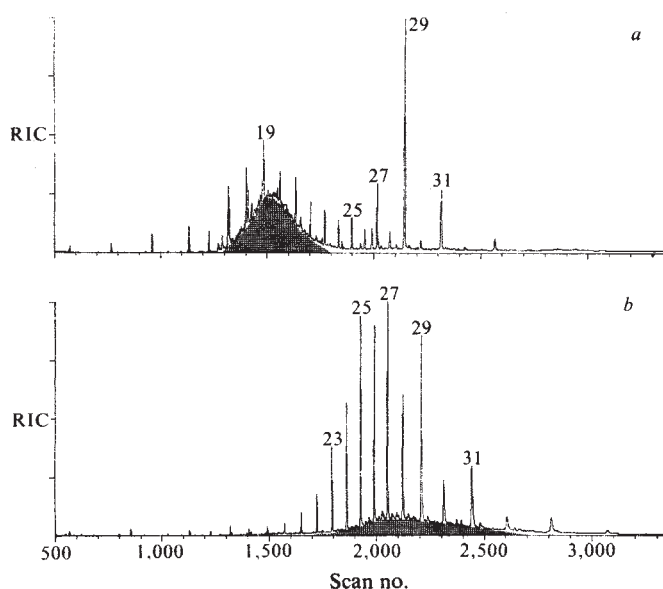


Fig. 2 Examples of total ion current traces from the GC-MS analyses of hydrocarbon fractions. (Numbers over the peaks refer to their respective carbon number.) a, Jos, upwind and b, Ibadan, urban.

Temperature of egg incubation determines sex in *Alligator mississippiensis*

Mark W. J. Ferguson

Anatomy Department, The Queen's University of Belfast, Medical Biology Centre, 97 Lisburn Road, Belfast BT9 7BL, UK

Ted Joanen

Louisiana Wildlife and Fisheries Commission, Rockefeller Wildlife Refuge, Grand Chenier, Louisiana 70643, USA

The factors controlling sexual differentiation in crocodilians are unknown, but heteromorphic sex chromosomes are absent from all species¹. Nichols and Chabreck² speculated that the sex of *Alligator mississippiensis* was not rigidly determined at the time of hatching but could be influenced by the post-hatching environment. They presented little evidence to support their hypothesis³ (no histological sections of hatchling gonads, no indication of the sex ratio at hatching), and their study failed to take account of habitat preferences of adult male and female alligators⁴. Here we demonstrate by laboratory and field experiments, that in *A. mississippiensis*: (1) Sex is fully determined at the time of hatching and naturally irreversible thereafter, and depends on the temperature of egg incubation, temperatures $\leq 30^\circ\text{C}$ producing all females, $\geq 34^\circ\text{C}$ yielding all males. (2) The temperature-sensitive period is between 7 and 21 days of incubation. (3) Natural nests constructed on levees are hotter (34°C) than those constructed on wet marsh (30°C), thus the former hatch males and the latter females. (4) The natural sex ratio at hatching is five females to 1 male. (5) Females hatched from eggs incubated at 30°C weigh significantly more than males hatched from eggs incubated at 34°C . This weight difference constitutes a possible selective evolutionary advantage of temperature-dependent sex determination (TSD) in alligators in that females become large and sexually mature as early as possible. The occurrence of TSD in alligators has wide-ranging implications for embryological, teratological, molecular, evolutionary, conservation and farming studies as well as for theories relating to the extinction of other Archosaurs.

Laboratory studies⁵⁻¹⁰ have indicated that in five families of turtles and two of lizards, the temperature of egg incubation affects the sex ratio of hatchlings. Only one study^{5,6} has investigated the effects of natural nest temperatures on sex determination in turtles, and even this involved the construction of artificial nests. There has been no demonstration of a selective evolutionary advantage for the occurrence of TSD in reptiles^{5,6}. Theoretically, such putative selective advantages could be considered within the framework of the Charnov-Bull model^{5,6,11,12}, which has not been experimentally confirmed in higher vertebrates. This model suggests that there is an advantage in allowing the sex of an embryo to respond to its immediate environment, that is, that hatchlings from a cold nest could become either good females or sub-standard males (or vice versa) and that the opposite is true for warm nests.

Alligator eggs (500) were removed from wild nests (Fig. 1) in the Louisiana swamps within 12 hours of egg laying. Only banded³ eggs, laid by middle-aged females, were used in laboratory experiments. These eggs have high fertility, a low spontaneous malformation rate and are homogeneous with respect to egg length, width and initial weight³ (see Table 2). Eggs were randomly placed in one of six groups and incubated (in nesting media³) at 26, 28, 30, 32, 34 or 36°C , in water-jacketed, air circulating incubators at 100% humidity (Table 1). Eggs incubated below 26°C or above 36°C died. After 60 days incubation

(the normal incubation period is 65 days), all eggs were killed and the enclosed alligators fixed in 10% formal saline. Using a dissecting microscope, we noted the presence or absence of oviducts or vas deferens, and the surface appearance of the gonads. The entire reproductive systems of all animals were excised, serially sectioned ($8\ \mu\text{m}$) in the transverse, sagittal and horizontal planes and stained with haematoxylin and eosin, periodic acid Schiff and haematoxylin, Mallory and Weigert Van Gieson. Hatchling gonads were histologically differentiated into ovaries or testes and showed no evidence of bisexuality or hermaphroditism. Sex is evidently fully determined by the time of hatching. Moreover, alligators raised for 1 year in a variety of conditions showed no signs of sex reversal and no evidence of bisexuality. In all experiments, sex was determined macroscopically as well as by serial histological sectioning.

All eggs incubated at $\leq 30^\circ\text{C}$ developed into females, whereas all those incubated at $\geq 34^\circ\text{C}$ developed into males (Table 1); eggs incubated at 32°C had a sex ratio of 86.7%:13.3% females to males. There was no significant difference in embryonic mortality between eggs incubated at 28°C , 30°C , 32°C or 34°C (Table 1). Even if all the dead embryos had been of opposite sex to those that were living,

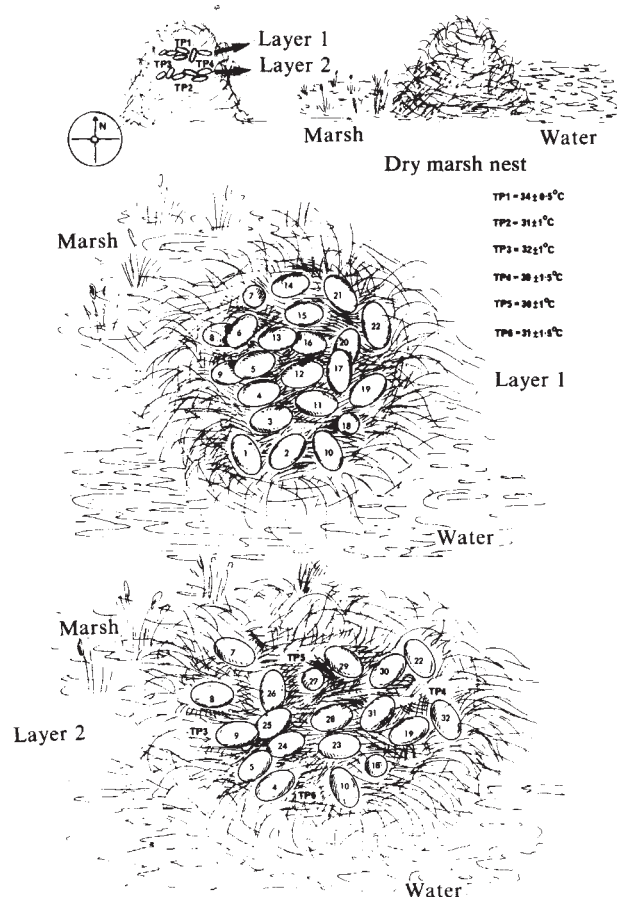


Fig. 1 Map showing the location of eggs and temperature probes TP1-TP6 in a dry marsh nest of *A. mississippiensis*. The nest consists of a mound of vegetation (mostly *Spartina patens*) and is ~4 ft in diameter and 3 ft high. Average temperature recorded by TP1 was $34 \pm 0.5^\circ\text{C}$, and for TP2-TP6, $31 \pm 1^\circ\text{C}$. Eggs 12, 13 and 16 contained male alligators (average intact weight of hatchlings = $37.5 \pm 0.5\text{ g}$, average weight of hatchling minus yolk = $28.8 \pm 0.3\text{ g}$, average weight of yolks = $8.5 \pm 0.5\text{ g}$); eggs 4-6, 8, 9, 14, 15, 20, 21, 25 and 27-31 contained female alligators (average intact weight of hatchlings = $39.1 \pm 2.8\text{ g}$, average weight of hatchling minus yolk = $28.7 \pm 4.4\text{ g}$, average weight of yolk = $10.7 \pm 1.4\text{ g}$); and eggs 1-3, 7, 10, 11, 17-19, 22-24, 26 and 32 were infertile, dead or malformed. The small size of these eggs indicated that they had been laid by a young female³.

Table 1 Sex and mortality of alligators removed after 60 days from eggs artificially incubated at varying temperatures

	Temperature of egg incubation ($\pm 0.2^\circ\text{C}$)					
	26 $^\circ\text{C}$	28 $^\circ\text{C}$	30 $^\circ\text{C}$	32 $^\circ\text{C}$	34 $^\circ\text{C}$	36 $^\circ\text{C}$
No. of eggs	50	100	100	100	100	50
No. of dead embryos (% of total)	40 (80)	4 (4)	3 (3)	2 (2)	6 (6)	43 (86)
Females % Of total eggs	20*	96*	97	85	0	0
% Of living eggs	100*	100*	100	86.7	0	0
Males % Of total eggs	0	0	0	13	94	14
% Of living eggs	0	0	0	13.3	100	100

All eggs were recovered from wild nests within 10 h of laying, and all were laid by middle-aged female alligators³. Infertile eggs were excluded and the average egg weights at the start of incubation were statistically similar in all temperature groups (Table 2).

* These animals were developmentally retarded and still exhibited yolk external to the body wall. They were at a stage of embryonic development equivalent to day 48 embryos incubated at 32 $^\circ\text{C}$. Thus they were excluded from the weighing analyses in Table 2.

the differences between the temperature groups would still be highly significant (Table 1). Thus, the temperature of egg incubation determines the sex of *A. mississippiensis* in laboratory experiments.

To determine the temperature-sensitive period during egg incubation critical for sex determination, eight groups each consisting of 20 eggs were incubated first at 30 $^\circ\text{C}$ and then switched to 34 $^\circ\text{C}$ after 1, 2, 3, 4, 5, 6, 7 and 8 weeks respectively. A comparable series commenced incubation at 34 $^\circ\text{C}$ and were switched to 30 $^\circ\text{C}$. We found that sex was determined by the incubation temperature during weeks 2 and 3; thereafter changes in incubation temperature had no effect. It was easier to produce females, particularly during week 3 of switches from 30 $^\circ\text{C}$ to 34 $^\circ\text{C}$. The temperature-sensitive period for alligators occurs much earlier than for other reptiles⁵⁻¹⁰, but alligator eggs are laid at an advanced embryonic stage³, so that the embryonic stages during which sex is determined are comparable. Again there was no evidence of sex reversal, rather the embryos showed progressive gonadal differentiation and were committed to a particular phenotype after the temperature-sensitive period.

Although we have shown that constant incubation temperatures determine sex in laboratory experiments, they may not do so in nature where nest temperatures fluctuate depending on the geographical location of the nest (levée, wet or dry marsh, shaded, exposed), the prevailing climate and whether embryos have a genetic disposition towards one sex^{5,6}. To investigate the role of temperature on sex determination in nature, we placed temperature probes (Taylor Instruments) connected to continuous chart recorders, in different locations (top, bottom and four sides of the egg cavity) within six typical wild alligator nests (see Fig. 1). Two nests were on dry leveés, two on wet marsh and two on dry marsh. The natural nest construction was not disturbed during insertion of the probes. Temperatures were recorded throughout the incubation period. After 60 days of incubation, nests were opened, the eggs counted and an accurate nest map drawn, to indicate the positions of the eggs and temperature probes (Fig. 1). All eggs were opened and the alligators coded, fixed and sexed. Sex of the animal was correlated with position in the nest and nest temperature. The average temperature of dry leveé nests was $34 \pm 0.7^\circ\text{C}$ at the periphery and bottom of the nests (TP2-TP6, Fig. 1) and $35 \pm 1^\circ\text{C}$ at the top centre of the nests (TP1, Fig. 1): 99% of living eggs ($n = 59$) were males. The average temperature of wet marsh nests was $29 \pm 1^\circ\text{C}$ at the periphery and bottom of the nests and $30 \pm 1^\circ\text{C}$ at the top centre of the nests: all living eggs ($n = 61$) were females. The average temperature of dry marsh nests was $31 \pm 1^\circ\text{C}$ at the periphery and bottom of the nests and $34 \pm 1^\circ\text{C}$ at the top centre (Fig. 1). Males hatched from eggs at the top centre of these nests and females from eggs at the periphery and bottom of the nests (Fig. 1). The average sex ratio for dry marsh nests was 5:1, females to males ($n = 63$). Evidently, sex is determined by egg incubation temperatures in nature as well as in laboratory experiments.

It has been argued^{5,6} that genotypic and environmental sex determination could co-exist if animals had a weak sex-determining locus which operated in most environmental conditions,

but which was overridden in extreme conditions. Predictions from Bull's model^{5,6} show that if genotype and temperature both determine sex, then a sex ratio of 1:1 should exist in the hatching population (but not necessarily in the adult population where differential sexual mortality may occur). To determine the sex ratio at hatching in a wild alligator population, we collected eggs (8,000) from all alligator nests laid within a large area (~2,000 acres) of typical habitat (including leveés and marsh) in the Rockefeller Wildlife Refuge, Louisiana. These eggs were collected after 5 weeks of incubation, after the temperature-sensitive period and before maximum egg predation by raccoons⁴. They were artificially incubated in nesting media³ at 32 $^\circ\text{C}$, allowed to hatch, raised in controlled environmental chambers and sexed externally⁴ at 1.5 yr old (we did not feel justified in killing 8,000 alligator hatchlings for microscopic sexing). As the sex ratio at hatching may vary from year to year, depending on climatic conditions, the above protocol was performed during 1978-81. The average sex ratio for the 4 years was $5 \pm 0.7:1$, females to males. This is consistent with field observations (marsh nests are more numerous than leveé nests) and with laboratory experiments, in which viable females are produced over a greater temperature range (28-32 $^\circ\text{C}$) than males (34 $^\circ\text{C}$). This heavily biased female sex ratio means that classical genotypic mechanisms probably have no role in alligator sex determination^{5,6}; but it also presents a problem for sex ratio theory. Fisher¹³ noted that selection within random-mating populations favours a primary sex ratio of 1:1, provided sons and daughters are equally costly to produce. Such a 1:1 result does not necessarily apply to TSD¹⁴, but even using Bull's theory, it is difficult to account for the extreme sex ratio observed in alligators. Therefore, either the present sex ratio is a recent development, perhaps resulting from changes in the availability of various nest sites, or some property of population structure or inheritance is peculiar and selects for the observed sex ratio.

We investigated possible selective advantages for the evolution of TSD in alligators. During incubation, yolk and albumen are progressively used for embryonic growth and metabolism¹⁵. Alligators hatch with a considerable volume of absorbed abdominal yolk which serves as a food source for the post-hatching period^{3,4}. As the rate of metabolism of poikilotherms, for example, alligators, depends on temperature¹⁶, we investigated whether there was any weight difference between male and female hatchlings incubated at 34, 32 or 30 $^\circ\text{C}$ both in the laboratory (Table 2) and in the field (Fig. 1). Since the size of any hatchling is related to the freshly laid egg weight³, the latter was controlled so that there were no significant differences between the temperature groups at the start of the experiment (Tables 1, 2). It was impossible, however, to control freshly laid egg weight in the nest map field experiments (Fig. 1), as this would have meant disrupting the nests. In laboratory experiments females hatched at 30 $^\circ\text{C}$ weighed significantly more than males hatched at 34 $^\circ\text{C}$ (Table 2). This weight difference is due to the females having significantly more absorbed yolk than the males (Table 2), which is consistent with the postulated lower rate of embryonic metabolism at 30 $^\circ\text{C}$. As all eggs were incubated at 100% humidity, the latter is unlikely

Table 2 Statistical analysis of the weights of day 60 alligator embryos (hatchlings) and their absorbed yolks as a function of egg incubation temperature and sex

Temperature of egg incubation ($\pm 0.2^\circ\text{C}$)	No. of eggs	Sex of hatchling	Initial egg weight (g)		Weight of intact hatchling (g)		Weight of hatchling minus yolk (g)		weight of yolk (g)	
30 °C	97	Female	65.1 $\pm$ 5.7	$P > 0.2$	47.6 $\pm$ 3.1	$P < 0.001$	37.8 $\pm$ 3.9	$P > 0.5$	8.1 $\pm$ 2.2	$P < 0.001$
34 °C	94	Male	66.2 $\pm$ 6.1		43.7 $\pm$ 4.0		38.2 $\pm$ 4.6		5.3 $\pm$ 1.9	
32 °C	85	Female	65.8 $\pm$ 6.8	$P > 0.7$	46.2 $\pm$ 4.5	$P > 0.7$	36.8 $\pm$ 4.2	$P > 0.8$	7.7 $\pm$ 2.5	$P > 0.9$
32 °C	13	Male			45.8 $\pm$ 3.9		37.0 $\pm$ 4.0		7.6 $\pm$ 2.8	

The initial egg weights were similar in all temperature groups, although in the 32 °C group it was clearly impossible to determine at the onset of the experiment which eggs would become males and which females. The hatchlings were fixed and weighed intact (minus extra-embryonic membranes) to the nearest 0.2 g. The absorbed yolk was then carefully dissected out of the abdomen and both it and the devoled hatchling weighed separately, to the nearest 0.2 g. Frequently the sum of the last two weights did not correspond exactly to the first weight due to very small losses of blood. Statistical comparison of any two means was calculated using the *t*-test¹⁸. Values are the mean $\pm$ s.d.

Table 3 Length and weight of 1-yr-old female and male alligators hatched from eggs incubated at 30 °C and 34 °C respectively

Sex	No. of animals	Total length (cm)	Weight (g)
Female (30 °C incubation)	12	63.25 $\pm$ 2.54	708.3 $\pm$ 78.5
Male (34 °C incubation)	4	53.34 $\pm$ 4.83	475 $\pm$ 108.5

The alligators were reared after hatching in artificial environmental chambers in identical conditions of, for example, temperature, food, light, water, exercise. Values shown are the mean $\pm$ s.d. The means for the two groups were compared using Student's *t*-test¹⁸. Total length was measured from the snout tip to the end of the tail.

to be important¹⁵, although decreased humidity could be associated with increased nest temperatures in the wild. Similar trends were observed in the field—levee nests produced lighter males, wet marsh nests heavier females. Moreover, in a single dry marsh nest, where the initial egg weights were similar, males hatched significantly lighter than females (Fig. 1).

The sexual difference in hatchling size and weight was maintained during the first year of life in alligators raised in environmental chambers⁴ in identical post-hatching conditions (Table 3). However, once alligators reach a length of ~1 m (at 2–3 yr old) the growth rate of females declines faster than that of males¹⁷. Thereafter, males grow faster and for a longer time than females¹⁷, probably as a result of their differing endocrinology. Sexual maturity in alligators depends on size, both sexes maturing when ~1.8 m long. Because of the initial heavy hatchling weight, and resultant rapid early growth, females become sexually mature either before, or at the same age, as males. The number of years of functional reproductive activity in females is about one-half that of males³. In addition, over 50% of the eggs laid by small (young) females are either infertile or malformed³, while social order favours breeding of larger females before smaller ones. These data, combined with the fact that one male can fertilize several females in any breeding season, make it highly desirable that female alligators should become large and sexually mature as quickly as possible. We suggest that this is a possible selective advantage for TSD in alligators. Thus, animals incubated at the low temperatures which produce females hatch with more energy reserves (absorbed yolk), develop, grow and become sexually mature faster than those incubated at the higher temperatures, which yield males (Tables 2, 3). If genotypic mechanisms determined sex, then the latter would be unrelated to incubation temperature and thus hatchling weight, so that 50% of the offspring would be light females. These animals would mature more slowly, enter the breeding population much later and contribute fewer offspring per lifetime than their heavier counterparts. In terms of the Charnov–Bull model^{11,12}, if the hatchling is light, there is a selective advantage in it being an average male, rather than a below average female. Temperature-dependent sex determination allows expression of this advantage by closely associating hatchling weight with sex. Depending on the life history of other reptiles, there may be selective advantages in being either a light or heavy female or male which could explain why low

temperatures produce females in lizards and males in turtles^{5–10}. Interestingly, if alligator eggs are incubated at 32 °C there is no difference in the weights of the resulting male and female hatchlings; both follow the heavier 30 °C female pattern (Table 2). Perhaps the heavy males hatching at this temperature become the large males in the adult population.

Male and female alligators occupy different micro-habitats after hatching: males inhabit open water canals beside levées and females the marsh⁴. Additional selective advantages, for example in terms of predation, may accrue from being a larger, faster growing female in the marsh or a smaller male in the open water. Interestingly, the habitat of adult male and female alligators is closely related to the geographical location of male- and female-producing nests (the former on levées beside open canals, the latter in the marsh). This may represent another selective advantage for TSD because neither male nor female hatchlings have to move great distances (when they would be subject to predation) from 'unsuitable' nesting sites to the preferred adult habitat.

By exploiting TSD to control the sex of alligator embryos experimentally, it is possible to study sexual differences in early embryogenesis as well as differences in the metabolism and action of teratogens³. In addition, TSD has wide implications for habitat planning in conservation studies and for pen design, egg collection times and artificial incubation temperatures in alligator farming. Moreover, if other ancient Archosaurs, for example, dinosaurs, also showed TSD with perhaps a different male–female threshold temperature, then this may explain the selective extinction of these groups in response to a relatively sudden, continuous change in climate to one that is either hotter or colder³.

We thank L. McNease and staff of the Rockefeller Wildlife Refuge for assistance with egg collection, J. Bull for valuable discussion and N. Steinberger for access to preliminary data, obtained by the late C. L. Yntema, which confirm our findings. This work was supported by MRC Grants G979/386/C and 8113610CB, grant EP109/74/75 from the Northern Ireland Eastern Health & Social Services Board, a research travel grant from the Wellcome Trust of Great Britain, a Dixon Research Scholarship from the Queen's University of Belfast, and a Churchill Fellowship, all awarded to M.W.J.F.

Received 16 October 1981; accepted 16 March 1982.

- Cohen, M. M. & Gans, C. *Cytogenetics* **9**, 81–105 (1970).
- Nichols, J. D. & Chabreck, R. H. *Am. Nat.* **116**, 125–137 (1980).
- Ferguson, M. W. J. *Archs oral Biol.* **26**, 427–443 (1981); *J. Craniofacial Genet. dev. Biol.* **1**, 123–144 (1981); *Trans. zool. Soc. Lond.* **36**, 99–152 (1982); in *Biology of the Reptilia* Vol. 12 (eds Gans, C. & Billett, F. S.) (in the press); in *The Reproductive Biology and Conservation of Crocodilians* (ed. Tryon, B. W.) (in the press); *Science* **214**, 1135–1137 (1981); *Evolution* (submitted).
- Joanen, T. & McNease, L. *Proc. A. Conf. SE Assoc. Fish & Wildl. Agencies* **23**, 141–151 (1969); **24**, 175–193 (1970); **25**, 106–116 (1971); **26**, 252–274 (1972); **28**, 482–500 (1974); **29**, 407–415 (1975); **32**, 179–181 (1978); *Int. Zoo Yb.* **19**, 61–66 (1979).
- Bull, J. J. & Vogt, R. C. *Science* **206**, 1186–1188 (1979).
- Bull, J. J. *Q. Rev. Biol.* **55**, 3–21 (1980).
- Pieau, C. *C. R. hebdom. Séanc. Acad. Sci., Paris* **272**, 3071–3074 (1971); **274**, 719–722 (1972); **277**, 2789–2792 (1973).
- Yntema, C. L. *J. Morph.* **125**, 219–252 (1968); **159**, 17–28 (1979).
- Yntema, C. L. & Mrosovsky, N. *Herpetologica* **36**, 33–36 (1980).
- Mrosovsky, N. *Am. Zool.* **20**, 531–547 (1980).
- Charnov, E. L. & Bull, J. J. *Nature* **266**, 828–830 (1977).

12. Charnov, E. L., Los-den Hartogh, R. L., Jones, W. T. & Van der Assem, J. *Nature* **289**, 27–33 (1981).
13. Fisher, R. A. *The Genetical Theory of Natural Selection* (Oxford, 1930).
14. Bull, J. J. *Heredity* **46**, 9–26 (1981).
15. Packard, G. C., Packard, M. J., Boardman, T. J. & Asher, M. D. *Science* **213**, 471–473 (1981).
16. Coulson, R. A. & Hernandez, T. *Biochemistry of the Alligator* (Louisiana State University Press, Baton Rouge, 1964).
17. Chabreck, R. H. & Joanen, T. *Herpetologica* **35**, 51–57 (1979).
18. Fisher, R. A. & Yates, F. *Statistical Tables for Biological, Agricultural and Medical Research* 6th edn (Longman, London, 1963).

Effect of *adh* genotype and heat stress on alcohol tolerance in *Drosophila melanogaster*

B. Sampsell & S. Sims

Department of Biological Sciences, Chicago State University,
Chicago, Illinois 60628, USA

The alcohol dehydrogenase locus (*adh*) is polymorphic in most natural populations of *Drosophila melanogaster*^{1–11}. Eight alleles have been identified on the basis of allozyme electrophoretic mobility and thermostability; two of these, *adh*^{Fm} and *adh*Sm, generally have combined frequencies of >90%. In four widely separate locations these major alleles display a clinal change in frequency, that of *adh*^{Fm} increasing with latitude and altitude^{3–6,11}. Biochemical studies have shown that the Fm allozyme (ADH^{Fm}) is more active, but less heat-resistant, than ADHSm (refs 12–14). The fact that these differences are associated with the geographical distribution of the alleles suggests that environmental temperature may be a selective factor in maintaining this polymorphism. We have investigated the effect of high temperature on flies having ADH allozymes with a range of thermostabilities and report here that flies which have relatively labile enzymes show reduced survival on alcohol-supplemented food compared with flies possessing more heat-resistant allozymes.

The specific physiological role of ADH has been much debated, but as the results of numerous experiments indicate that alcohol tolerance is a function of overall ADH activity levels^{15–20}, the enzyme is thought to contribute to the flies' ability to feed and breed in fermenting fruits. The relatively low temperature at which ADH can be inactivated *in vitro* (40–43 °C)^{5,7,14} suggests that a similar effect could be produced in living flies encountering these temperatures, for example while feeding on fallen fruits warmed by the sun or by heat of fermentation. As ADH is necessary for survival on alcoholic media, alcohol tolerance should be reduced if a significant portion of a fly's ADH is denatured by non-lethal heat exposure. Under such heat and alcohol stress, the relative thermostabilities of different allozymes may have adaptive significance.

In designing an experiment to test this hypothesis, it was desirable to eliminate, as far as possible, the effects of other loci which might alter ADH activity. We therefore used a breeding scheme designed to derive five strains, each homozygous for a different *adh* allele, but with nearly uniform genetic backgrounds. By using laboratory stocks with various visible mutations and cross-over suppressors, strains were developed which differ only in a small region (about six map units long) surrounding the *adh* locus (II, 50.1). The strains all have an unmarked X chromosome derived from the Cy/Pm, D/Sb marker stock (stock 97; Amherst, Massachusetts); second chromosomes in which both ends (from 0 to 48.5⁺ and from 51.5⁺ to 108) came from marker stock P232 (Pasadena, California); and third chromosomes obtained from a male caught in the wild. The fourth chromosome was uncontrolled. A sixth stock was not involved in the breeding scheme; this stock was obtained from Davis, California (stock R 72) and carries *adh*^{Fs}

(strain 1 in Fig. 1). Allozyme and allele symbols are based on the electrophoretic mobility and thermostability of the enzyme: Fr, fast mobility, heat-resistant; Fm, fast mobility, moderately heat-sensitive; Fs, fast mobility, heat-sensitive; Sm, slow mobility, moderately heat-resistant; Ss, slow mobility, heat-sensitive.

Male flies from uncrowded culture vials were collected soon after eclosion and aged for 5 days on unsupplemented instant *Drosophila* medium at 23 °C. For the heat treatments, flies were placed in empty, cotton-stoppered shell vials containing a piece of moist blotting paper to maintain adequate humidity. Vials were heated for 13 min in a 40 °C water bath. The flies were then allowed to recover in the same vials for ~4 h by which time they were standing or walking around. The treatment period of 13 min was chosen as the longest time for which almost all flies could survive. It was observed that very small flies usually did not recover consciousness. This occurred sporadically in all the strains and did not seem to be genotype-specific. After the recovery period, flies were transferred to vials of instant medium supplemented with various concentrations (by volume) of ethanol. After 2 days the number of live flies was counted to determine percentage survival. Control flies were treated in the same way, but were not heated. Figure 1 compares the survival of heated with untreated flies of each strain; differences between strains are apparent.

The alcohol tolerance of the heat-treated flies was also compared with that of the control flies for each of the six strains by examining the slopes of the regression lines obtained for each treatment (Table 1). The untreated control flies showed no difference among strains in their response to alcohol. Although there was a slight reduction in survival for both 6% and 8% alcohol, this reduction could not be related to variations in the ADH thermostabilities. By contrast, alcohol tolerance was clearly decreased after heat treatment, and the three strains having the most sensitive enzymes, ADH^{Ss} and ADH^{Fs}, showed the greatest reduction. That the reduced survival on ethanol was a result of a decrease in ADH activity was confirmed by direct measurements of ADH activity levels in heat-treated flies (Table 2). Spectrophotometric assays revealed that flies with ADH^{Ss} and ADH^{Fs} lose nearly 90% of their ADH activity when subjected to 40 °C for 13 min. Flies with ADH^{Fm} have about one-third less ADH, but as these flies had the highest initial levels of ADH activity, a substantial amount of enzyme remains effective. Finally, flies with the two most resistant allozymes, ADH^{Fr} and ADHSm, have the same levels of enzyme activity after heat exposure as before. Figure 1 shows that *adh*^{Fr} flies have the same alcohol tolerance regardless of heat stress, in the range of ethanol concentrations used. In the case of *adh*Sm, there was a significant reduction in survival following heat treatments, but the slope of the regression line was not significantly different from that for the controls, indicating that survival on ethanol was not interacting with the heat treatment.

The 95% confidence intervals of the slopes of the regression lines for the heat-treated flies overlap to such an extent that the only strains that are significantly different are *adh*^{Fr} and *adh*Sm, compared with the two *adh*^{Fs} strains. Note, however,

Table 1 Slopes of regression lines for each strain and treatment in Fig. 1

Strain	Control		Heat-treated	
	Slope	95% Confidence interval	Slope	95% Confidence interval
<i>adh</i> ^{Fr}	-0.45	-1.48–0.58	-0.35	-1.43–0.73
<i>adh</i> Sm	-1.50	-5.00–2.00	-2.05	-4.25–0.15
<i>adh</i> ^{Fm}	-1.10	-1.90–0.30	-3.10	-5.51–0.70
<i>adh</i> ^{Ss}	-0.70	-1.22–0.18	-5.40	-10.49–0.31
<i>adh</i> ^{Fs}	0.00		-7.50	-10.50–4.50
(strain 1) <i>adh</i> ^{Fs}	-1.10	-1.60–0.60	-9.75	-12.33–7.17
(strain 2)				

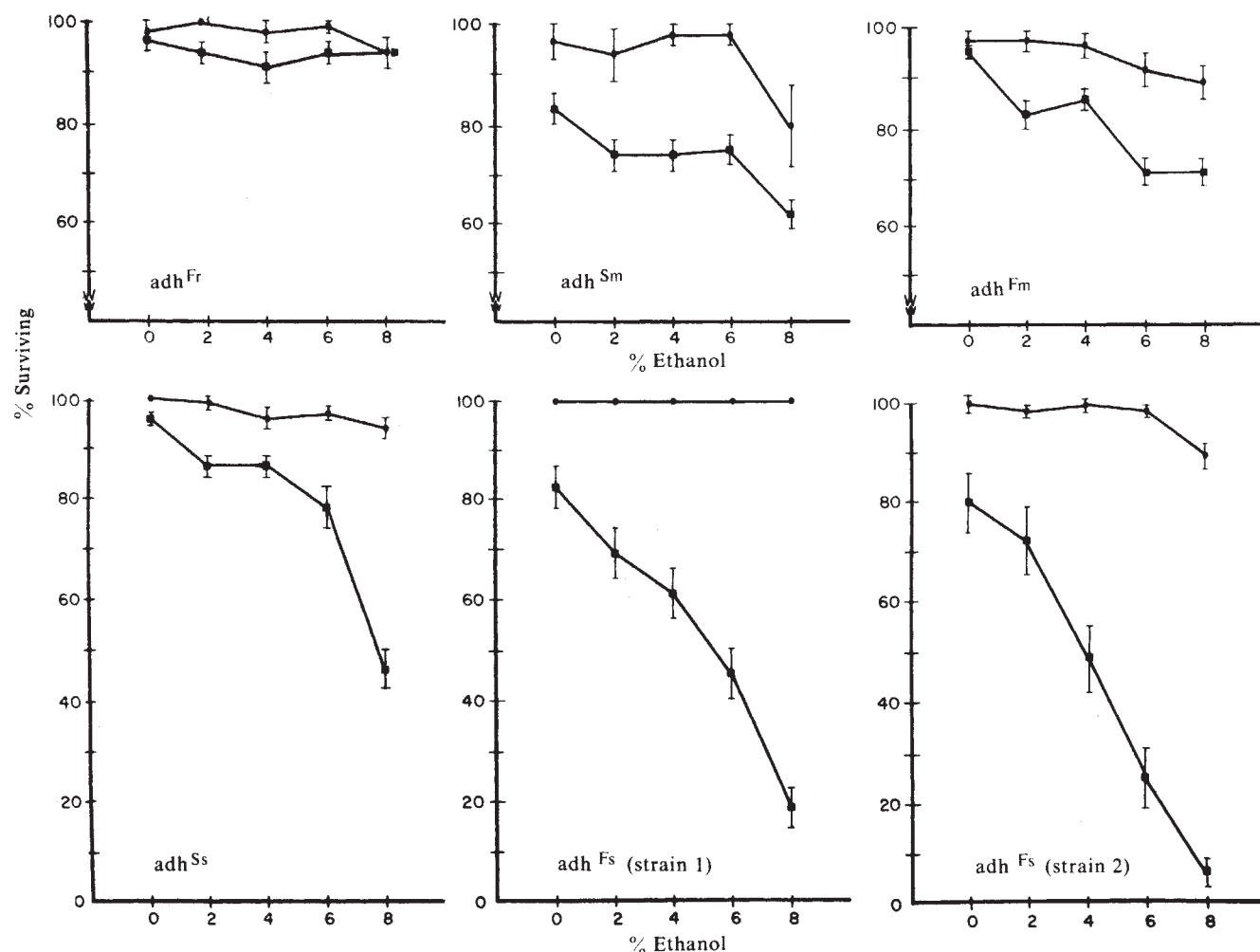


Fig. 1 Relative survival of untreated (control, ●) and heat-treated (40 °C, 13 min, ■) flies on various concentrations of ethanol. Each data point represents the survival of 100–800 flies. Bars indicate 95% confidence intervals calculated from the binomial distribution.

that if the strains are ordered by their slopes (indicating interaction of the heat and alcohol stress) the order obtained is exactly the same as the *in vitro* decrease in allozyme thermostability previously reported: $Fr > Sm > Fm > Ss > Fs$ (ref. 21).

Our observations suggest that the variant alleles, *adh^{Ss}* and *adh^{Fs}*, may be rare in nature because they are conditionally deleterious or even lethal. Fitness of flies carrying these alleles depends on environmental temperature combined with alcohol concentration.

Other investigators have examined the survival of flies with different *adh* genotypes following heat shock. Johnson and Powell²² tested flies from four natural populations. In two of the samples, they observed a higher frequency of *adhSm* among the survivors of a heat shock (40 °C for 30 min) than among a group of untreated flies. No increase in the frequency of either allele was observed in the other two samples. Milkman²³ found no difference in *adh* frequency between the survivors and non-survivors after heat treatment (40 °C for 10 min). It should be emphasized that Johnson and Powell did not expose their flies, either during or after the heat treatments, to any situation in which ADH is known to be essential.

In our experiments, heat-treated flies with different allozymes showed similar levels of survival in the absence of additional ethanol: ~95% for *adh^{Fr}*, *adh^{Fm}* and *adh^{Ss}* strains, and 82% for *adhSm* and *adh^{Fs}* strains. Only when the food was supplemented with at least 4% ethanol did the survival differences among the strains become significantly correlated with enzyme thermostability. Thus the relevance of Johnson and Powell's results²² to the maintenance of the *adh* polymorphism is not clear.

A correlation between environmental temperature and *adh* allele frequency has been observed for the major alleles in Spain⁹, along the eastern coast of the United States²⁴ and in southern Texas and eastern Mexico^{4,6}. The relationship between temperature and fitness however, may not be a simple one. In a more thorough analysis of the north-south clinal variation in a series of Australasian populations, Anderson²⁵ examined the relationship between *adh* frequency and various environmental variables that change with latitude. He found that differences

Table 2 ADH activity in heat-treated and untreated control flies

Strain	ADH activity (units per mg wet weight per min)		% Reduction
	Untreated controls	Heat-treated (40 °C, 13 min)	
<i>adh^{Fr}</i>	16.6 (0.3)	17.1 (0.4)	None
<i>adh^{Fm}</i>	29.8 (1.3)	20.2 (0.1)	32
<i>adh^{Fs}</i>	19.4 (1.0)	2.6 (0.4)	87
<i>adhSm</i>	9.8 (0.1)	10.0 (0.2)	None
<i>adh^{Ss}</i>	12.3 (1.5)	1.5 (0.2)	88

ADH activity was measured at 25 °C in a 2 ml reaction mixture containing 2.7 M ethanol and 1 mM NAD in 100 mM Tris-HCl buffer pH 8.6, plus 0.2 ml of crude extract. The extracts were prepared by homogenizing 30 male flies in 1 ml of Tris buffer and centrifuging at 10,000 r.p.m. for 30 min. One unit of ADH activity represents an increase in absorbance of 0.001 at 340 nm. Activities represent the average for two separate groups of flies; standard deviations are shown in parentheses.

in rainfall accounted for 76% of the variation in allele frequency. None of the other variables, including several temperature statistics, could explain the remainder of the variation. He then suggested that temperature and humidity could influence indirectly the *adh* polymorphism through their effects on such environmental factors as competitor and predator species, types of vegetation available as food sources, varieties of decay-causing bacteria and yeasts, and rates of fermentation and evaporation of metabolic products including ethanol.

Received 16 November 1981; accepted 8 March 1982.

- O'Brien, S. J. & MacIntyre, R. J. *Am. Nat.* **103**, 97–113 (1969).
- Berger, E. M. *Genetics* **66**, 677–683 (1970).
- Grossman, A. I., Koreneva, L. G. & Ulitskaya, L. E. *Soviet Genet.* **6**, 211–214 (1970).
- Pipkin, S. B., Rhodes, C. & Williams, N. J. *Hered.* **64**, 181–185 (1973).
- Vigue, C. L. & Johnson, F. M. *Biochem. Genet.* **9**, 214–227 (1973).
- Pipkin, S. B., Franklin-Springer, E., Law, S. & Lubega, S. J. *Hered.* **67**, 258–266 (1976).
- Sampsel, B. *Biochem. Genet.* **15**, 971–987 (1977).
- Voelker, R. A., Mukai, T. & Johnson, F. M. *Genetica* **2**, 143–148 (1977).
- Malpica, J. M. & Vassallo, J. M. *Nature* **286**, 407–408 (1980).
- Wilkes, A. V., Gibson, J. B., Oakeshott, J. G. & Chambers, G. K. *Aust. J. biol. Sci.* **33**, 575–585 (1980).
- McKay, F. T. C. *Genet. Res.* **37**, 79–93 (1981).
- Gibson, J. B. & Miklovich, R. *Experientia* **27**, 99–100 (1971).

It is almost certain that selection acting on the *adh* locus will involve a complex set of factors and mechanisms. Our results indicate that temperature can, in appropriate conditions, act directly on fitness through the inactivation of heat-sensitive enzymes.

This research was supported by NSF grant DEB 79-11538 and NIH grant RR-08043-10. Some equipment was purchased through NIH training grant T32-GM07091. We thank R. Milkman and G. Eertmoed for reading the manuscript and W. Prigge for assistance with graphics.

- Day, T. H., Hillier, P. C. & Clarke, B. *Biochem. Genet.* **11**, 141–165 (1974).
- Thorig, G. E. W., Schoone, A. A. & Scharloo, W. *Biochem. Genet.* **13**, 721–731 (1975).
- Briscoe, D. A., Robertson, A. & Malpica, J. M. *Nature* **255**, 148–149 (1975).
- Morgan, P. *Heredity* **34**, 124–127 (1975).
- McDonald, J. F. & Avise, J. C. *Biochem. Genet.* **14**, 347–355 (1976).
- McDonald, J. F., Chambers, G. K., David, J. & Ayala, F. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4562–4566 (1977).
- Thompson, J. N. & Kaiser, T. N. *Heredity* **38**, 191–195 (1977).
- Kamping, A. & van Delden, W. *Biochem. Genet.* **16**, 541–551 (1978).
- Sampsel, B. M. & Milkman, R. *Biochem. Genet.* **16**, 1139–1141 (1978).
- Johnson, F. M. & Powell, A. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1783–1784 (1974).
- Milkman, R. *Drosoph. Inf. Serv.* **52**, 58 (1977).
- Johnson, F. M. & Schaffer, H. E. *Biochem. Genet.* **10**, 149–163 (1973).
- Anderson, P. R. in *Genetic Studies of Drosophila Populations* (eds Gibson, J. B. & Oakeshott, J. G.) 237–250 (Australian National University Press, Canberra, 1981).

Signal processing times in bacterial chemotaxis

Jeffrey E. Segall, Michael D. Manson* & Howard C. Berg

Division of Biology 216–76, California Institute of Technology, Pasadena, California 91125, USA

The bacterium *Escherichia coli* responds to changes in the concentrations of various chemicals in its environment¹. A cell swims along a smooth trajectory (runs), moves erratically for a brief time (tumbles) and then runs again, choosing a new direction at random². If a run happens to carry the cell up a gradient of an attractant (such as aspartate, serine and certain sugars), the occupancy of the appropriate chemoreceptor increases with time^{3,4} and a signal is sent to the flagellar motors that increases their counterclockwise bias⁵. On the average, this extends favourable runs and the cell moves up the gradient. The receptors for aspartate and serine^{6–8} are proteins found in the cytoplasmic membrane, known as methyl-accepting chemotaxis proteins^{9,10}, and are the products of the *tar* and *tsr* genes¹¹. A cell can adapt to sustained changes of receptor occupancy by carboxymethylating these proteins¹²; it is not known, however, how these proteins signal the flagellar motors or how the signal controls the direction of flagellar rotation. Products of several *che* genes involved in signalling and adaptation have been identified^{13,14}, but with the exception of a methyltransferase¹⁵ (the *cheR* product) and a demethylase¹⁶ (the *cheB* product), their functions are largely unknown. In an attempt to learn more about the events that trigger a chemotactic response, we have now exposed cells to rapid changes in the concentration of attractants and repellents and measured the time required for flagellar reversal. In wild-type cells and in cells containing a *cheR-cheB* deletion, the response latency is ~0.2 s. In *cheZ* mutants, it is much longer.

Cells were tethered to glass with antibodies directed against flagellar filaments, so that rotation of a cell body could be used to monitor the rotation of a flagellar motor¹⁷. Individual cells were observed with a microscope equipped with an opto-electronic device that provided a continuous readout of the cell's angular position^{18,19}. A micropipette containing sodium α -methyl-DL-aspartate (α -MeAsp, a non-metabolizable ana-

logue of the attractant aspartate) was positioned with its tip a few micrometres away from the cell, and the cell was stimulated iontophoretically by the passage of current. As the addition of α -MeAsp causes a cell to spin counterclockwise⁵, the current was switched on when the cell changed its direction of rotation from counterclockwise (CCW) to clockwise (CW). We measured the interval of time from the beginning of this stimulus to the next reversal (the response latency), with a resolution of ~0.03 s. A typical response is shown in Fig. 1. Stimuli were given about every 30 s, and distributions of interval lengths were constructed (see Fig. 2). The experiments were repeated with different concentrations of α -MeAsp in the pipette and the results are summarized in Table 1. When the concentration in the pipette was ≤ 0.01 mM, the distributions of interval lengths did not differ significantly from those observed in the absence of a stimulus (Fig. 2a). When the concentration was ≥ 0.1 mM, responses occurred in ~0.2 s (Fig. 2b; Table 1). When the concentration was increased by a factor of 100, the mean latency decreased by less than a factor of 2 (Table 1).

To interpret these results, one needs to know the initial concentration of α -MeAsp near a cell and the time course of the change generated by the stimulus. A cell exposed to a step change in concentration of α -MeAsp spins CCW for a time (the transition time) proportional to the change in receptor occupancy. This time is ~180 s for a shift from 0 to 0.16 mM, the apparent dissociation constant of the receptor^{20,21}. The initial concentration of α -MeAsp near a cell was estimated by suddenly bringing the pipette from far away to its final position. No response was detected with ≤ 1 mM α -MeAsp in the pipette. With 10 mM α -MeAsp in the pipette, a transition time of ~20 s was observed, corresponding to a jump in concentration at the cell from 0 to ~0.008 mM. The maximum concentration change that could be generated by the stimulus was estimated by measuring the transition time after the pipette was in position and the current was switched on and left on. Steady-state concentrations determined in this way are given in the last column of Table 1. Negative ions carried out of the pipette by the current ultimately reach the cell by diffusion. For diffusion from a point source switched on at time $t=0$ (ref. 22), the concentration at the cell increases as $C=C_{\infty} \text{ERFC}[r/(4Dt)^{1/2}]$, where C_{∞} is the steady-state concentration, ERFC is the error function complement, r is the distance between the cell and the tip of the pipette ($\sim 4 \times 10^{-4}$ cm) and D is the diffusion constant of α -MeAsp (8.9×10^{-6} cm² s⁻¹). The concentration rises rapidly, reaching one-third, one-half and two-thirds of its steady-state value in 0.01, 0.02 and 0.05 s

* Present address: Fakultät für Biologie, Universität Konstanz, D-7750 Konstanz 1, FRG.

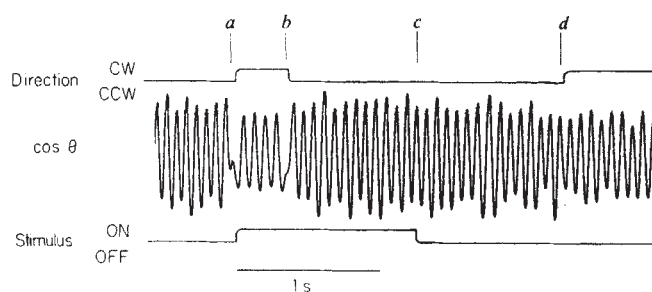


Fig. 1 A strip-chart record of iontophoretic stimulation of tethered *E. coli* strain AW405. The middle trace shows the cosine of the angle of orientation of the cell, $\cos \theta$; $\sin \theta$ also was recorded (not shown). These two signals were used to generate the event marker shown in the upper trace, which indicates the direction of rotation of the cell. The event marker was updated every time $\sin \theta = 0$, that is, twice per cycle, and was used to trigger the stimulus. A second event marker, shown in the lower trace, was turned on when current was injected into the pipette²². The latency was determined by measuring the interval between the onset of the stimulus and the next reversal, as indicated by $\sin \theta$ or $\cos \theta$. The measurement was corrected for the lag in the electronics used to smooth these signals (0.06 ± 0.02 s). A spontaneous reversal occurred at *a*. The cell was stimulated 0.03 s after this event was flagged by the upper event marker. Another reversal (the response) occurred at *b*, 0.34 s later (0.28 s when corrected for lag in the electronics). The stimulus ended at *c*, and the cell changed direction again at *d*. Cells were grown as described previously²⁰ on glycerol rather than glucose, washed twice with a tethering medium containing 90 mM NaCl, 10 mM KCl, 10 mM Tris pH 7.0 and 0.1 mM tetraethylenepentamine (tetren³³), resuspended in this medium at one-tenth the original volume, sheared by passage between two syringes equipped with 26-gauge needles and washed once more. Then they were tethered, as described previously²⁰. The micropipette (F. Haer 30-31-1 Pyrex 7740 tubing pulled on a D. Kopf 700C puller) was filled with tethering medium containing, in addition, 10 mM sodium α -MeAsp (Sigma) and 0.01 mM ThCl_4 (see below). Its resistance was about 50 M Ω . Connection to the current-injection circuit was made with AgCl-coated Ag wires. The tip of the pipette was positioned 2 μm above a point 4 μm away from the point of attachment of the cell. The stimulus current was -100 nA (negative ions ejected). No positive back current was used. The bath was perfused slowly with tethering medium to prevent any long-term buildup of attractant. The cell was observed with a $\times 40$ water-immersion phase-contrast objective (Zeiss) electrically insulated from the microscope. The experiment was done at room temperature ($22 \pm 1^\circ\text{C}$). ThCl_4 was added to the medium in the pipette to neutralize negative charges on the surface of the glass³⁴ which we think were responsible for stimulus artefacts (responses that occurred without any α -MeAsp in the pipette, particularly to the passage of positive currents). As thorium is precipitated by phosphate and chelated by EDTA, these components of the standard tethering medium²⁰ were replaced by Tris and tetren respectively.

respectively. With 0.1 mM α -MeAsp in the pipette, the cells responded on the average in 0.32 s (Table 1), at which time the receptor occupancy was 2.3%. With 1 mM or 10 mM α -MeAsp in the pipette, this change in receptor occupancy should have occurred in 0.002 s or 0.004 s, respectively. The cells responded on the average 0.25 or 0.23 s later, when the changes in receptor occupancy were 10 or 69%, respectively. We conclude that the cells respond with a latency of ~ 0.2 s when the change in receptor occupancy is as small as 2% and that larger changes in receptor occupancy do not change the latency much. The latency is well within the upper limit for a chemotactic response time (~ 0.7 s) set by the rapid-mixing experiments of Macnab and Koshland³.

The proper functioning of the pipettes (see ref. 23) was checked by replacing the α -MeAsp with 10 mM fluorescein and measuring the fluorescence near the tip of the pipette with a photomultiplier mounted on a dark-field microscope. Changes in fluorescence occurred during the passage of a pulse of negative current on a time scale consistent with that predicted by the diffusion theory to within ~ 0.01 s.

We looked for correlations between the mean latency of a cell and the pre-stimulus directional bias of its flagellar motor. In the 34 cells examined, the ranges of mean CW intervals, mean CCW intervals and fractions of time spent CW were 0.34–1.89 s, 0.66–3.12 s and 0.11–0.71, respectively. The mean latencies were not correlated with any of these parameters.

The following controls demonstrated that we were dealing with *bona fide* chemotactic responses. Wild-type cells (strain AW405) did not respond when the pipette contained only tethering medium (Table 1); they responded to benzoate (100 mM in the pipette) as a repellent, that is, by spinning CW in response to a pulse of negative current. A *tar* mutant (strain AW539) failed to respond to α -MeAsp but did respond to benzoate as a repellent. A *tsr* mutant (strain AW518) responded to both α -MeAsp and benzoate as attractants, as expected²⁴. Whenever a response occurred, the latency was ~ 0.2 s.

What steps in the signal-processing pathway cause the latency? Recall that the latency changes very little when the concentration of α -MeAsp is increased by a factor of 100 (Table 1). With large stimuli, more attractant can diffuse to the cell, enter the periplasm and bind to the receptor in a given time, so none of these processes can be limiting. Reactions whose rates are proportional to the amount of bound receptor cannot limit the latency either, because the limit is reached with changes in receptor occupancy as small as 2%. The times required for a small molecule to diffuse from the receptor to the flagellar motor or for an electrogenic impulse to travel along the cytoplasmic membrane are a few milli- or microseconds respectively, so signal transmission *per se* is not likely to be limiting. Finally, the time required for the flagellar motor to change the direction of rotation of a tethered cell once a reversal is initiated is 0.01 s or less²⁵, so this factor is not important. Either the *tar* protein activates a sequence of one or more reactions that signal the motor with a delay averaging 0.2 s, or the motor, having received the signal, is not able to initiate a reversal in a shorter time.

The latency of a strain containing a *cheR-cheB* deletion is normal (strain RP1273)²⁶. Therefore, excitation does not involve methylation¹², demethylation²⁶ or other processes catalysed by the *cheR* or *cheB* gene products. The latencies of strains containing *cheZ* mutations are very long, ~ 2 s (strains RP5006 = *cheZ* 292, an amber mutant, and RP5007 = *cheZ* 293, ref. 27). It is known from reversion analysis and interspecies complementation tests that the *cheZ* gene product interacts with both the demethylase (the *cheB* gene product) and with components of the flagellar motor involved in controlling the direction of rotation (the *cheC* = *flaA* and *cheV* = *flaB*

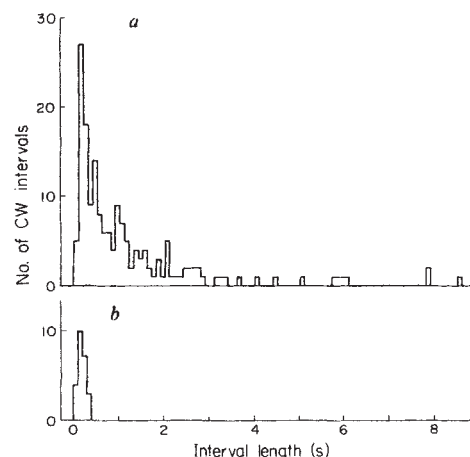


Fig. 2 Distributions of rotation intervals observed before (*a*) and after (*b*) iontophoretic stimulation of *E. coli* strain AW405 (as indicated in Fig. 1). Distribution *a* includes CW intervals observed immediately preceding each stimulus (3–5 intervals per stimulus, 171 in all, mean $\pm$ s.d. 1.32 ± 1.93 s). Distribution *b* includes CW intervals observed during each stimulus (1 interval per stimulus, 24 in all, mean $\pm$ s.d. 0.18 ± 0.10 s, corrected for the lag in the electronics). Distribution *a* is exponential, except for a paucity of intervals of length ≤ 0.1 s; distribution *b* is bell-shaped. The distribution of CCW intervals observed immediately preceding each stimulus (not shown, 179 intervals in all, mean $\pm$ s.d. 1.36 ± 1.40 s) was similar to the corresponding distribution of CW intervals (*a*).

Table 1 Response of the wild-type strain to different concentrations of α -MeAsp

Concentration in pipette (mM)	CW interval (s)		No. of cells	Total no. of stimuli	Steady-state concentration at the cell (mM)*
	Before stimulus	During stimulus			
10	0.89 $\pm$ 0.31	0.23 $\pm$ 0.07	8	100	0.52
1	0.89 $\pm$ 0.50	0.25 $\pm$ 0.05	9	173	0.022
0.1	0.90 $\pm$ 0.47	0.32 $\pm$ 0.09	10	164	0.0044
0.01	1.18 $\pm$ 0.34	0.92 $\pm$ 0.23	3	26	<0.0005
0	1.04 $\pm$ 0.10	1.00 $\pm$ 0.49	3	31	0

Measurements were made as described in Figs 1 and 2 legends, with -100 nA pulses of length 0.3–5 s. Each cell was tested at only one concentration of α -MeAsp in the pipette. Values for the CW interval are the mean and s.d. of the mean intervals for each cell, each cell weighted equally.

* Estimated from the transition time, t , for adaptation to a continuous -100 nA pulse, using the formula $C = 0.16 t / (360 - t)$, where C is in mM, t in s (see Fig. 2 of ref. 20). The fractional change in receptor occupancy is $C / (K_d + C)$, where K_d is the apparent dissociation constant of the receptor.

gene products)²⁸. Our results suggest that the *cheZ* product lies on the excitation pathway. Alternatively, its binding to the flagellar motor is required for rapid initiation of CCW rotation.

The lack of correlation between latency and the directional bias of the motor observed in wild-type cells also seems to be the case for cells containing *cheR-cheB* or *cheZ* mutations. The time required for a cell to adapt to a step change in the concentration of α -MeAsp, however, is strongly correlated with the directional bias of the motor (with the mean CCW interval)²⁹. This implies that the signals involved in excitation and adaptation reach the motor by separate pathways or that the latency is due to a step that occurs after these pathways merge.

Although *cheZ* mutants generally have abnormally high tumbling rates, *cheZ-cheC* double mutants exist that fail to move up gradients even though their tumbling rates are normal³⁰. Mutants of this sort may represent a new kind of *che* phenotype, in which the defect lies in the rates at which cells respond to time-varying stimuli. If the latency of a cell is longer than its mean run length, it will choose a new direction at random before it has had a chance to make the appropriate response.

Why is the latency of wild-type cells as long as 0.2 s? There is no reason, *a priori*, that the response to a large stimulus could not occur within a few milliseconds. However, a cell must time-average its measurement of receptor occupancy if it is to sense small changes in concentration in the presence of statistical fluctuations. For most of the stimuli encountered in nature, averaging times of hundreds of milliseconds are required³¹. There may never have been any selective pressure for the development of more rapid signal-processing machinery.

We thank Steve Block, Jim Hudspeth, Victor Neher and Bob Smyth for valuable suggestions and Sandy Parkinson for bacterial strains. This work was supported by grant AI16478 from the US NIH.

Received 14 December 1981; accepted 15 March 1982.

- Adler, J. *Science* **166**, 1588–1597 (1969).
- Berg, H. C. & Brown, D. A. *Nature* **239**, 500–504 (1972).
- Macnab, R. & Koshland, D. E. *Jr Proc. natn. Acad. Sci. U.S.A.* **69**, 2509–2512 (1972).
- Brown, D. A. & Berg, H. C. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1388–1392 (1974).
- Larsen, S. H., Reader, R. W., Kort, E. N., Tso, W.-W. & Adler, J. *Nature* **249**, 74–77 (1974).
- Clarke, S. & Koshland, D. E. *Jr J. biol. Chem.* **254**, 9695–9702 (1979).
- Hedblom, M. L. & Adler, J. *J. Bact.* **144**, 1048–1060 (1980).
- Wang, E. A. & Koshland, D. E. *Jr Proc. natn. Acad. Sci. U.S.A.* **77**, 7157–7161 (1980).
- Kort, E. N., Goy, M. F., Larsen, S. H. & Adler, J. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3939–3943 (1975).
- Springer, M. S., Goy, M. F. & Adler, J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3312–3316 (1977).
- Silverman, M. & Simon, M. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3317–3321 (1977).
- Springer, M. S., Goy, M. F. & Adler, J. *Nature* **280**, 279–284 (1979).
- Parkinson, J. S. A. *Rev. Genet.* **11**, 397–414 (1977).
- Silverman, M. & Simon, M. I. A. *Rev. Microbiol.* **31**, 397–419 (1977).
- Springer, W. R. & Koshland, D. E. *Jr Proc. natn. Acad. Sci. U.S.A.* **74**, 533–537 (1977).
- Stock, J. B. & Koshland, D. E. *Jr Proc. natn. Acad. Sci. U.S.A.* **75**, 3659–3663 (1978).
- Silverman, M. & Simon, M. *Nature* **249**, 73–74 (1974).
- Kobayashi, S., Maeda, K. & Imae, Y. *Rev. scient. Instrum.* **48**, 407–410 (1977).
- Berg, H. C., Manson, M. D. & Conley, M. P. *Symp. Soc. exp. Biol.* **35**, 1–31 (1981).
- Berg, H. C. & Tedesco, P. M. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3235–3239 (1975).
- Spudich, J. L. & Koshland, D. E. *Jr Proc. natn. Acad. Sci. U.S.A.* **72**, 710–713 (1975).
- Crank, J. *The Mathematics of Diffusion*, 2nd edn (Clarendon, Oxford, 1975).
- Purves, R. D. *J. Neurosci. Meth.* **1**, 165–178 (1979).

- Muskavitch, M. A., Kort, E. N., Springer, M. S., Goy, M. F. & Adler, J. *Science* **201**, 63–65 (1978).
- Berg, H. C. *Cold Spring Harb. Conf. Cell Proliferation* **3**, A47–A56 (1976).
- Sherris, D. & Parkinson, J. S. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6051–6055 (1981).
- Parkinson, J. S. *J. Bact.* **135**, 45–53 (1978).
- Parkinson, J. S. *Symp. Soc. gen. Microbiol.* **31**, 265–290 (1981).
- Spudich, J. L. & Koshland, D. E. *Jr Nature* **262**, 467–471 (1976).
- Parkinson, J. S. & Parker, S. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2390–2394 (1979).
- Berg, H. C. & Purcell, E. M. *Biophys. J.* **20**, 193–219 (1977).
- Dreyer, F. & Peper, K. *Pflügers Arch. ges. Physiol.* **348**, 263–272 (1974).
- Reilley, C. N. & Vavoulis, A. *Analyt. Chem.* **31**, 243–248 (1959).
- Schanne, O. F., Lavalée, M., Laprade, R. & Gagné, S. *Proc. IEEE* **56**, 1072–1082 (1968).

Low doses of ethanol disrupt sensory responses of brain noradrenergic neurones

Gary Aston-Jones, Stephen L. Foote & Floyd E. Bloom

A. V. Davis Center for Behavioral Neurobiology, The Salk Institute, PO Box 85800, San Diego, California 92138, USA

Ethanol is a widely abused drug which has many behavioural and psychological effects¹. In spite of considerable research^{1–4}, the brain mechanisms responsible for these effects are unknown. Previously, it has been proposed that noradrenaline (NA)-containing locus coeruleus (LC) neurones, which project throughout the brain^{5,6}, mediate the effects of many abused as well as clinically effective psychoactive agents^{7,8}. Recent studies^{9–11} have shown that in freely behaving, undrugged animals, NA-containing LC (NA-LC) neurones exhibit marked, short-latency responses to sensory stimuli of many modalities, perhaps serving to bias brain and behavioural activities towards adaptive responses to phasic, unexpected environmental events. We have now examined the effects of ethanol on these sensory responses of NA-LC neurones. In the present study, anaesthetized animals were used to minimize fluctuations in arousal, providing a more stable baseline for assessing pharmacological effects. A class of NA-LC sensory responses which mimic those observed in unanaesthetized animals was studied. In addition, using antidromic stimulation, we investigated the effects of ethanol on the soma excitability, axonal conduction velocity, and strength of recurrent collateral inhibition of these neurones. We now report that low intoxicating doses of ethanol substantially reduce the magnitude and temporal reliability of sensory-evoked responses in NA-LC neurones, perhaps due in part to enhanced feedback inhibition of these cells.

Action potentials were recorded from single NA-LC neurones using glass micropipettes in anaesthetized, male adult albino rats. Rectal temperature was maintained between 36

and 37 °C using a thermistor-controlled heating pad. The locus coeruleus was approached at a 30° caudo-rostral angle to avoid the overlying transverse sinus. The experimental paradigm consisted of recording at least 3 min of spontaneous discharge, followed by activity during a 1 Hz train of 100 antidromic stimuli to either the medial or lateral neocortex, or to the dorsal noradrenergic bundle. Thirty second later activity was recorded

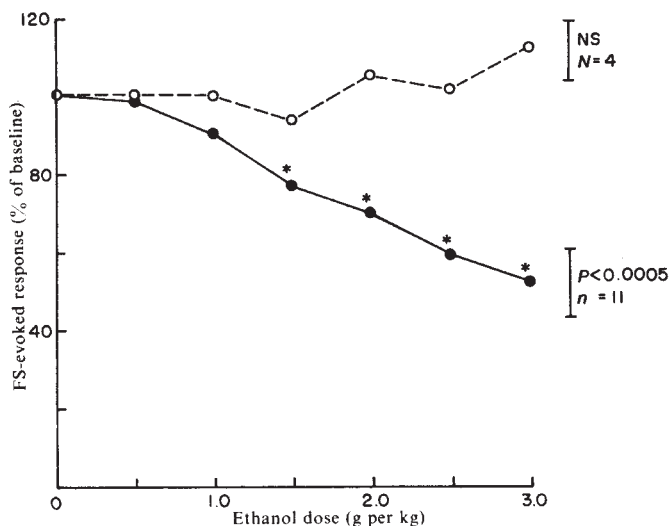


Fig. 1 Effects of ethanol on the magnitudes of FS-evoked response. Response magnitudes are plotted for cumulative doses of ethanol (●) as per cent of baseline (that is, before ethanol treatment) values. Data for similar consecutive trains of FS stimulation without ethanol administration are also shown (○). Data from animals anaesthetized with chloral hydrate ($n=5$) and halothane ($n=6$) are pooled. NS, no statistically significant effect of stimulation alone on response magnitudes. Bars represent mean $\pm$ s.e.m. and P values are confidence levels resulting from one-way analysis of variance over ethanol doses. * $P<0.005$, paired t -tests for each dose compared with baseline.

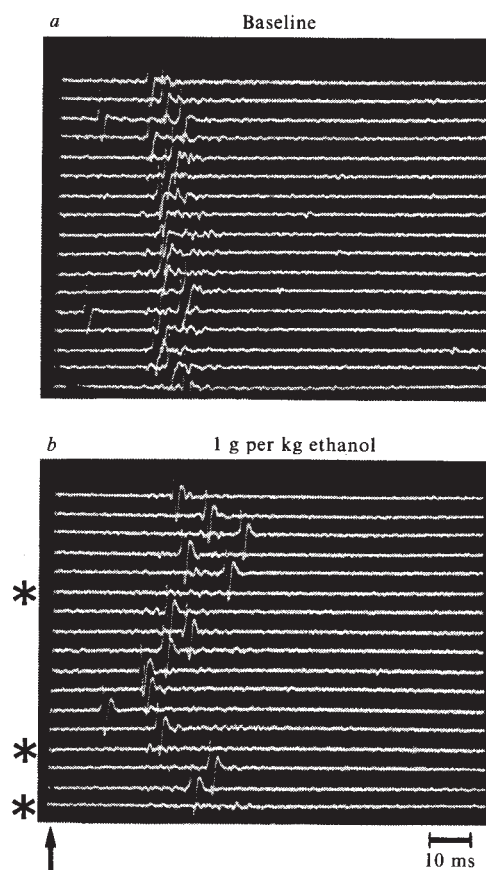


Fig. 2 Effect of ethanol on the latencies of FS-driven spikes. Oscilloscope sweeps triggered consecutively from top to bottom by 1 Hz FS stimuli (arrow) before (a) and 10 min after 1 g per kg ethanol i.p. (b) for the same NA-LC neurone. Note the three sweeps in b that have no spikes (indicated by asterisks) and the increased variability in the latency of FS-driven spikes compared with pre-ethanol activity in a.

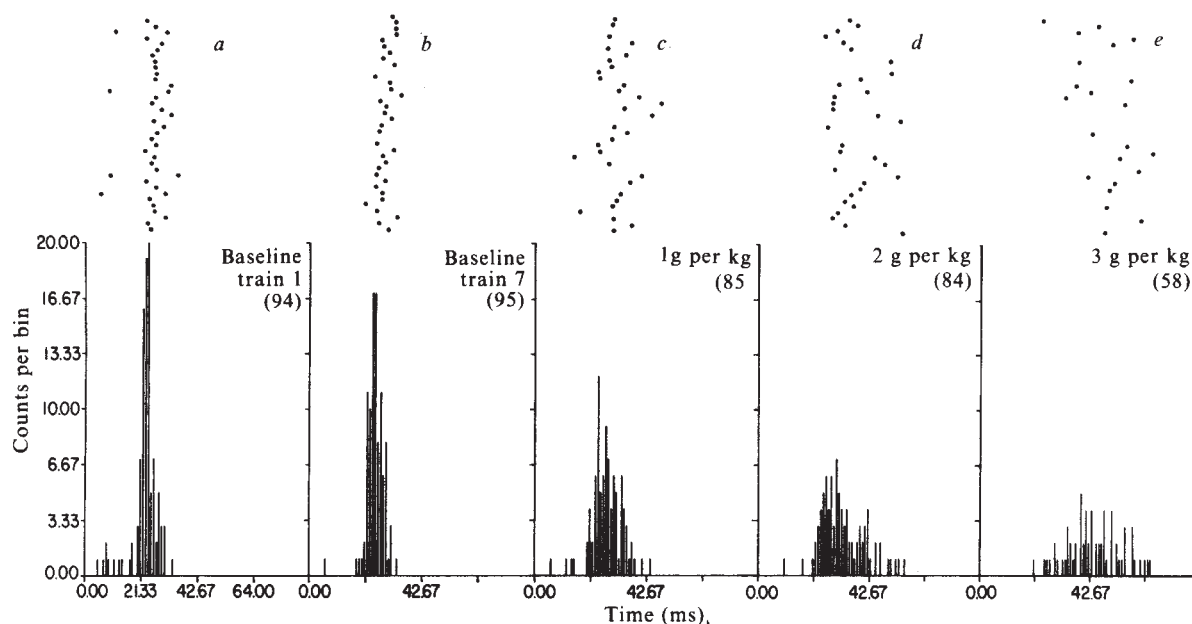


Fig. 3 Post-stimulus time histograms (PSTHs) and raster displays of activity in the same FS-stimulated NA-LC neurone before and after administration of cumulative ethanol doses. a, b, PSTHs for the first and seventh trains of 100 pre-ethanol stimuli and corresponding raster displays for 35 of these stimuli (in consecutive order from top to bottom). c-e, Data following cumulative doses of ethanol, as shown. Numbers in parentheses are the number of FS-driven spikes occurring in each corresponding PSTH. Note that the increased latency variability of FS-driven activity can occur without substantial change in the magnitudes of driven responses (compare c with d). PSTH time bases also apply to corresponding raster displays.

during a similar train at 10 Hz and a subsequent 1 Hz train of 100 electrical stimuli to the contralateral foot (foot-shock, FS). Responses were quantified using computer-generated post-stimulus time histograms (PSTHs)⁹. Response magnitude was defined as the number of spikes within a specified PSTH response interval. After determination of baseline values, 0.5–1.0 g ethanol (12% aqueous solution) per kg was injected intraperitoneally (i.p.), and then the series of tests was repeated. Ethanol injections and test series were made at 10 min intervals, up to a maximum cumulative dose of 3 g per kg. Only one cell per animal was tested with ethanol, and blood ethanol levels were verified at the end of the recording sessions to be appropriate for the injected dose. All recordings were histologically localized to the LC by iontophoresing Pontamine Sky Blue from the tip of the micropipette at the recording site.

Ethanol had no statistically significant effect on the mean spontaneous discharge rate of NA-LC neurones across the entire dose range tested in animals anaesthetized with either chloral hydrate or halothane. Similarly, ethanol had no effect on the ability of 1 Hz antidromic stimulation to elicit soma impulses in these neurones. Pronounced effects of ethanol were seen, however, on the ability of electrical FS stimulation to drive NA-LC neurones orthodromically. As shown in Fig. 1, ethanol decreased in a dose-dependent manner the magnitudes of responses elicited by FS stimuli, an effect observed in each NA-LC neurone tested. Response magnitudes were significantly reduced by a lower dose of ethanol (1 g per kg) in chloral hydrate- than in halothane-anaesthetized (1.5 g per kg) animals. The decrease in number of FS-driven spikes was the result of neither repeated trains of FS stimulation nor prolonged anaesthesia (see Fig. 1).

In addition to reducing response magnitudes, ethanol increased substantially the variability of trial-to-trial response latencies exhibited by individual cells (Figs 2, 3). This may be a more profound effect of ethanol on NA-LC neurones than the decrease in response magnitudes, because the standard deviation of response latencies increased to a much greater extent from baseline values (325% after 3 g per kg; see Fig. 4) than the response magnitudes decreased (55% below baseline after 3 g per kg ethanol). In addition, the increase in the standard deviation of latencies was significantly different from baseline at a lower dose of ethanol than was the decrease in response magnitude (compare Figs 1, 4). Thus, the temporal reliability of responses was reduced in a dose-dependent manner for NA-LC neurones after the administration of ethanol, but more markedly and at lower doses than the parallel decreases in response magnitudes. Furthermore, these results for single cells indicate that the temporal coherence (that is, synchronous response activity) characteristic of NA-LC neurones⁹ may be markedly reduced by these low doses of ethanol.

A mechanism by which ethanol may exert these effects on NA-LC neurones is suggested by its effects on antidromically elicited activity. Although ethanol had no effect on the number of soma spikes elicited by 1 Hz antidromic stimulation, antidromic soma activation by 10 Hz stimulation and the spontaneous activity which follows antidromic invasion were both reduced by ethanol in a similar dose-dependent manner, to ~60 and 70% of baseline, respectively, after 3 g per kg ethanol ($P < 0.01$ and < 0.05 , respectively; one-way analysis of variance, $n = 5$ for all measurements). As both these effects are indices of potency of the recurrent collateral inhibition previously proposed for NA-LC neurones¹², the results indicate that ethanol may increase such recurrent inhibition following activity within the NA-LC. This increased inhibition may partially explain the reduced sensory responsivity we found for these neurones.

In addition we found that ethanol caused a significant dose-dependent increase of >8% in the latency of antidromic initial segment spikes in NA-LC neurones during 1 Hz activation. This increased lability in conduction velocity along NA-LC axons seems to enhance the normal variability of conduction

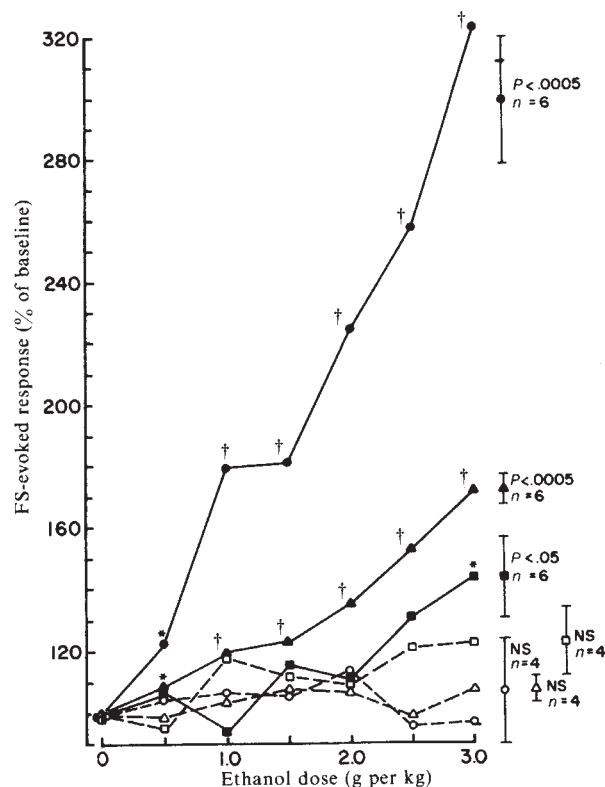


Fig. 4 Effect of cumulative ethanol doses on various measures of FS-driven activity in single NA-LC neurones. FS-response latency values: ●, s.d. for ethanol treatment; ○, s.d. for control data (obtained for similar trains of FS stimulation in the absence of ethanol); ▲, mean for ethanol; △, mean for control; ■, onset for ethanol; □, onset for control. Data are expressed as per cent of baseline values. Note that the most substantial effect of ethanol is to increase the variability (expressed here as s.d.) of latency to FS-driven activity, and that there is only minimal effect on onset latencies. Bars represent mean $\pm$ s.e.m. and P values are confidence levels resulting from one-way analysis of variance for dose of ethanol, or for number of stimulus trains in the case of control data. NS, not statistically significant. * $P < 0.05$; † $P < 0.005$, paired t -tests for dose or stimulus train compared with baseline data.

velocity during low-frequency impulse activity previously reported for these neurones¹³.

Our results showing decreased sensory responsiveness of NA-LC neurones after low doses of ethanol suggest a new interpretation of previous reports of the effects of ethanol on NA-LC discharge. In one study in paralysed rats¹⁴, most LC neurones exhibited decreased activity after a 2 g per kg dose. However, the cells were found to discharge spontaneously at ~17 Hz, much more rapidly than is typical of NA-LC neurones in active animals^{9–11} or in anaesthetized rats^{12,13,15–19}. If this elevated activity is due to increased tonic orthodromic input to NA-LC neurones (perhaps as a result of stress induced in paralysed, artificially respired, unanaesthetized animals), then the decreased activity following ethanol administration would be consistent with our results of reduced sensory responsivity in these cells after ethanol. This interpretation agrees with previous results²⁰ in anaesthetized rats, in which more characteristic rates of spontaneous activity for NA-LC neurones were unaffected by ethanol.

In the present studies, quantitatively similar effects of ethanol on NA-LC discharge were observed using either chloral hydrate or halothane anaesthesia; similar results were also observed for three animals anaesthetized with urethane. The fact that these results are reproducible with different anaesthetics indicates that these ethanol effects are not due to an interaction of ethanol

with the anaesthetic and that similar effects may be observed in waking preparations. The fact that doses of ethanol (for example, 3 g per kg) that are greater than those yielding significant reductions in sensory responsiveness in our studies had no effect on the ability of 1 Hz antidromic spikes to invade the soma-dendritic membrane indicate that our results are not due to a general membrane effect on these neurones, such as decreased excitability of NA-LC somata. In addition, previous studies²¹ indicate that effects of ethanol on peripheral nerve conduction are unlikely to account for our results. The lack of effect on spontaneous discharge rate provides additional evidence that ethanol at these doses is not acting in a nonspecific manner, and also indicates that the increase in intra-coerulear collateral inhibition we propose is not sufficiently intense to alter discharge during spontaneous, unsynchronized activity. Rather, this effect appears to become manifest only when collaterals are synchronously activated, such as during antidromic or sensory stimulation when a large percentage of NA-LC neurones would be engaged simultaneously.

In conclusion, our results demonstrate that low intoxicating doses (0.5–1.0 g per kg) of ethanol produce a marked decrease in the magnitude of sensory-evoked activity in NA-LC neurones and substantially increased variability in the latency of these responses, while also increasing the lability of impulse conduction along their axonal projections. Our results further lead us to hypothesize that ethanol increases the potency of recurrent collateral inhibition within the NA-LC, an effect perhaps responsible in part for the reduced sensory responsiveness of these neurones. We cannot exclude, however, the possibility that ethanol may also decrease the efficacy of excitatory afferents to the locus coeruleus.

Our results suggest a possible mechanism for many of the actions of ethanol on the brain. The diverse efferent system of projections originating from NA-LC neurones^{5,6} and the pronounced effects of NA on locus coeruleus target cells^{22–25} indicate that a change in the discharge characteristics of NA-LC neurones would have widespread consequences on the central nervous system and on behaviour. Previous studies on the normal discharge properties of NA-LC neurones in behaving animals^{9–11} have suggested that this system functions to bias the mode of brain information processing and the orientation of behaviour to favour either phasically adaptive responses to unexpected external stimuli or more tonic, vegetative behavioural programmes^{26,27}. Significant disorganization of such a function, an implied action of ethanol in the present studies, may underlie the marked disruption by ethanol of adaptive behavioural activity.

Conduction block in the peripheral nervous system in experimental allergic encephalomyelitis

M. P. Pender & T. A. Sears

Sobell Department of Neurophysiology, Institute of Neurology, University of London, Queen Square, London WC1N 3BG, UK

Experimental allergic encephalomyelitis (EAE) has been widely studied as a model of multiple sclerosis, a central nervous system (CNS) disease of unknown aetiology. The clinical features of both EAE and multiple sclerosis provide the only guide to the progress and severity of these diseases, and are used to assess the response to treatment. In such comparisons the clinical features of EAE are assumed to be due to lesions in the CNS, but in this disease there is also histological evidence of damage to the peripheral nervous system^{1–8}. However, the functional consequences of such peripheral lesions have been entirely ignored. To examine this we have studied nerve conduction in rabbits with EAE. We report here that most of the large diameter afferent fibres are blocked in the region of the dorsal root ganglion and at the dorsal root entry zone, thus accounting for the loss of tendon jerks and also, through the severe loss of proprioceptive information, the ataxia of these animals. We conclude that whenever clinical comparisons are made between EAE and multiple sclerosis, the pathophysiology associated with the histological damage of the peripheral nervous system must be taken into account.

EAE is a disease induced by inoculating with CNS-derived antigens and adjuvants, and is characterized clinically by limb ataxia and weakness⁹ and histologically by CNS lesions consisting of meningeal infiltration with mononuclear cells, perivascular cuffing and para-adventitial infiltration with mononuclear cells and perivascular demyelination^{9,10}. This disease has been studied extensively, especially recently in its chronic relapsing form¹¹, because it is widely accepted as an animal model of multiple sclerosis¹². However, it has been assumed that the clinical features of EAE are due to CNS lesions even though histological damage of the peripheral nervous system occurs in many species such as the rabbit^{3,4}, mouse⁵, guinea pig^{1,5} and monkey² and when antigens derived solely from the CNS are used³. It also occurs in rabbits with chronic EAE⁶, and in guinea pigs⁷ and rats⁸ with chronic relapsing EAE. Thus, it is possible that the clinical features of EAE are due to lesions of the peripheral rather than the central nervous system. This question needs resolving because the improvement or suppression of the clinical features of EAE by agents such as myelin basic protein¹³ and immunosuppressants¹⁴ provides the basis for their use in the treatment of multiple sclerosis.

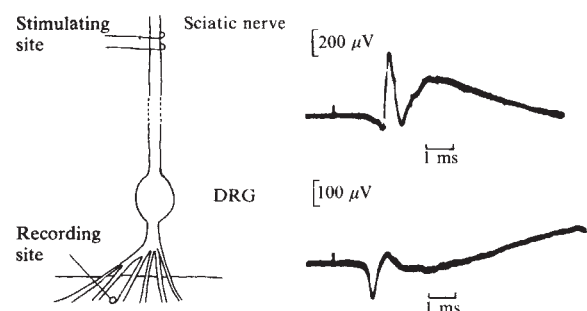


Fig. 1 Recording in the volume conductor over the S1 dorsal root entry zone in response to sciatic nerve stimulation. In each trace, positivity at the active electrode gives a downward deflection. Upper trace, normal rabbit; lower trace, rabbit with EAE. Note the difference in gain. DRG, dorsal root ganglion.

Received 12 November 1981; accepted 17 February 1982.

- Sandler, M. (ed.) *Psychopharmacology of Alcohol* (Raven 1980).
- Siggins, G. R. & Bloom, F. E. *Pharmac. Biochem. Behav.* **13**, Suppl. 1, 203–211 (1980).
- Rogers, J., Siggins, G. R., Schulman, J. A. & Bloom, F. E. *Brain Res.* **196**, 183–198 (1980).
- Newlin, S. A., Mancillas-Trevino, J. & Bloom, F. E. *Brain Res.* **209**, 113–128 (1981).
- Ungerstedt, U. *Acta physiol. scand. Suppl.* **367**, 1–48 (1971).
- Pickel, V., Segal, M. & Bloom, F. E. *J. comp. Neurol.* **155**, 15–42 (1974).
- Schildkraut, J. J. & Kety, S. S. *Science* **156**, 21–30 (1967).
- Aghajanian, G. *Nature* **276**, 186–188 (1978).
- Aston-Jones, G. & Bloom, F. E. *J. Neurosci.* **1**, 887–900 (1981).
- Aston-Jones, G. & Bloom, F. E. *J. Neurosci.* **1**, 876–886 (1981).
- Foot, S. L., Aston-Jones, G. & Bloom, F. E. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3033–3037 (1980).
- Aghajanian, G., Cedarbaum, J. & Wang, R. *Brain Res.* **136**, 570–577 (1977).
- Aston-Jones, G., Segal, M. & Bloom, F. E. *Brain Res.* **195**, 215–222 (1980).
- Pohorecky, L. A. & Brick, J. *Brain Res.* **131**, 174–179 (1977).
- Nyback, H., Walters, J., Aghajanian, G. & Roth, R. *Eur. J. Pharmac.* **32**, 302–312 (1975).
- Svensson, T., Bunney, B. & Aghajanian, G. *Brain Res.* **92**, 291–306 (1975).
- Falens, A. & Mogenson, G. *Expl Neurol.* **53**, 254–266 (1976).
- Nakamura, S. & Iwama, K. *Brain Res.* **99**, 372–376 (1975).
- Takigawa, M. & Mogenson, G. *Brain Res.* **135**, 217–230 (1977).
- Svensson, T. & Engberg, G. in *Biological Effects of Alcohol* (ed. Regleiter, H.) 535–550 (Plenum, New York 1980).
- Kalant, H. *Fedn Proc.* **34**, 1930–1941 (1975).
- Foot, S. L., Freedman, R. & Oliver, A. P. *Brain Res.* **86**, 229–242 (1975).
- Segal, M. & Bloom, F. E. *Brain Res.* **72**, 79–97 (1974).
- Segal, M. & Bloom, F. E. *Brain Res.* **72**, 99–114 (1974).
- Woodward, D. J., Moises, H. C., Waterhouse, B. D., Hofer, B. J. & Freedman, R. *Fedn Proc.* **38**, 2109–2116 (1979).
- Aston-Jones, G. thesis, California Institute of Technology (1981).
- Aston-Jones, G., Foot, S. L. & Bloom, F. E. in *Norepinephrine: Clinical Aspects* (eds Ziegler, M. & Lake, C.) (Williams and Wilkins, Baltimore, in the press).

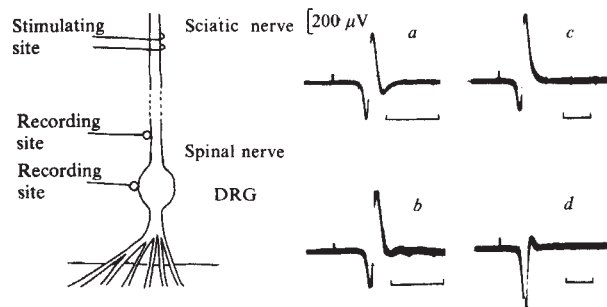


Fig. 2 Recordings in the volume conductor over the spinal nerve and the DRG in response to sciatic nerve stimulation. *a*, Recording over S1 spinal nerve in normal rabbit. *b*, Recording over normal S1 DRG. *c*, Recording over S1 spinal nerve of rabbit with EAE. *d*, Recording over S1 DRG from the same rabbit with EAE. Time bars, 2 ms.

We have therefore made a clinical, physiological and histological study of EAE in New Zealand white rabbits inoculated intradermally with homogenized rabbit spinal cord and complete Freund's adjuvant. The clinical onset occurred 15–23 days after inoculation and was typical of that previously described⁹, with weight loss, ataxia and weakness of the limbs occurring over 3–4 days. We also examined the tendon jerks and found that these usually disappeared about 2 days after the clinical onset, indicating a lesion interrupting the monosynaptic reflex arc. All the clinical findings could thus be accounted for by lesions in the peripheral nervous system. Histological examination revealed typical, widespread CNS lesions as well as lesions in the dorsal root ganglia and dorsal and ventral roots.

In terminal experiments, 2–6 days after clinical onset, animals were anaesthetized intravenously with 25% urethane, a lumbo-sacral laminectomy was performed and the left sciatic nerve was exposed in the posterior thigh. The left medial gastrocnemius compound muscle action potential elicited by stimulation of the sciatic nerve in the mid-thigh was normal in all of 12 animals studied.

Conduction into the roots and spinal cord was studied in 10 affected animals. Figure 1 shows typical responses recorded in the volume conductor over the left S1 dorsal root entry zone in response to sciatic nerve stimulation. The upper trace (from a normal control animal) shows a triphasic (positive-negative-positive) wave representing the afferent volley, followed by a longer latency, slow negative wave, the N wave. The initial positive wave is due to the passive outward currents driven by the last activated nodes of Ranvier; the negative phase of the afferent volley is due to the active inward currents occurring during the rising phase of the action potentials and ranges from 300 to 800 μ V in amplitude. The peak-to-peak amplitude of the triphasic wave ranges from 400 to 1,000 μ V. The N wave is due to the synaptic currents in the second order dorsal horn neurones excited mainly by low threshold cutaneous afferents.

The lower trace in Fig. 1 (from an affected animal) shows three characteristic abnormalities. (1) A very reduced peak-to-peak amplitude (usually <200 μ V) indicating that only a small proportion of the large diameter afferent fibres are conducting past the dorsal root ganglion to this point. (2) A very reduced (usually <50 μ V) negative wave and a relative increase in the positive wave of the afferent volley, indicating that most of the large diameter fibres still conducting past the ganglion are blocking in the vicinity of the dorsal root entry zone. (3) A delayed peak of the N wave possibly due to conduction persisting in small fibres.

To determine the site of conduction block, volume conductor recordings were made over the spinal nerve and dorsal root ganglion (Fig. 2); traces *a* and *b* are the responses over the normal S1 spinal nerve and S1 dorsal root ganglion, respectively. They are similar and consist of an initial positive wave followed by a larger negative wave. Figure 2*c* shows a recording

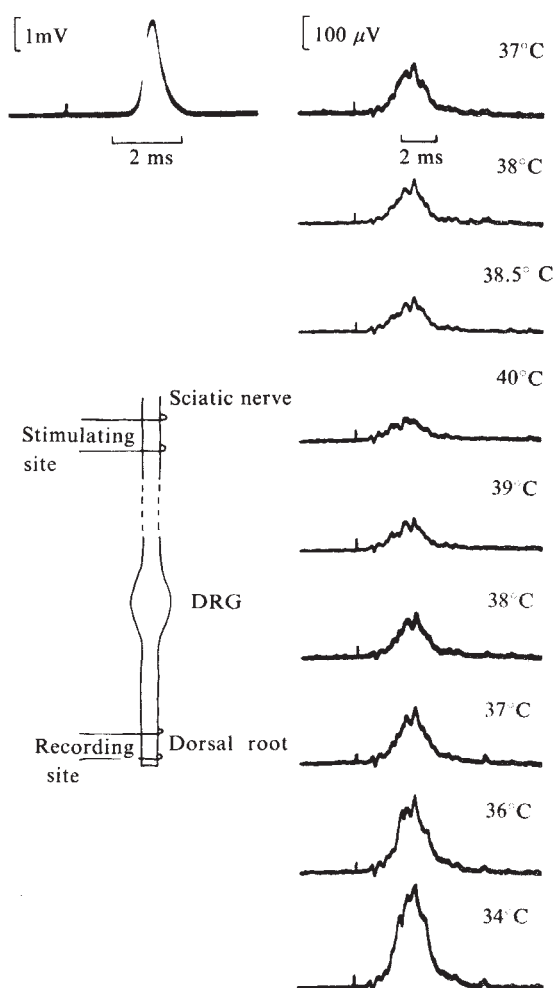


Fig. 3 Monophasic recording from the distal cut end of S1 dorsal root in response to sciatic nerve stimulation. Upper left trace, normal rabbit. Traces on the right are recordings from a rabbit with EAE at various laminectomy pool temperatures, as indicated on the right. Note the difference in gain and time base.

over the S1 spinal nerve of a rabbit with EAE and is normal. The recording in Fig. 2*d* was obtained over the S1 dorsal root ganglion of the same affected animal and shows a large positive wave followed by a small negative wave, indicating that most of the fibres are blocking at this point.

To investigate the nature of the conduction block, the effects of temperature were studied. Recently, it has been shown^{15,16} that in demyelinated single fibres, increasing the temperature within the physiological range produces reversible conduction block, and that in some fibres conduction block can be overcome by reducing the temperature by as little as 0.5 °C. The method used to study the effects of temperature is shown diagrammatically in Fig. 3. The distal cut end of the S1 dorsal root was placed on a pair of electrodes and a monophasic recording made of the compound action potential in response to sciatic nerve stimulation. The upper left trace is from a normal anaesthetized rabbit. The right-hand traces were made at various temperatures of the laminectomy oil pool of a rabbit on the fifth day of clinical disease. On the previous day we increased the body temperature of this animal by 1.5 °C—this caused a reversible increase in ataxia. The response at 37 °C was greatly reduced in amplitude (note the difference in gain) with an area less than 30% of the normal compound action potential, suggesting that most of the large diameter fibres were blocking in

the region of the dorsal root ganglion. The area of the response was progressively reduced as the temperature was increased to 40 °C, which indicates that block was occurring in an increasing number of fibres. When the temperature was reduced to 37 °C, the response regained its original area. On further cooling to 34 °C the area increased by 100%, showing that conduction was restored in a significant number of fibres. These findings strongly suggest that the conduction block is due to demyelination.

From our studies we conclude that the loss of tendon jerks and the ataxia of rabbits with EAE can be accounted for by conduction block in most of the large diameter afferent fibres in the region of the dorsal root ganglion and dorsal root entry zone. Such conduction block would mask the clinical expression of any central lesions which alone could produce ataxia. Clearly, these functional considerations must be taken into account when making a clinical comparison between EAE and multiple sclerosis.

We thank Drs L. Cuzner and C. Bolton for advice concerning the inoculation procedure. Support from the National Fund for Research into Crippling Diseases and the Brain Research Trust is gratefully acknowledged.

Received 18 January; accepted 9 March 1982.

- Freund, J., Stern, E. R. & Pisani, T. M. *J. Immun.* **57**, 179–194 (1947).
- Ferraro, A. & Roizin, L. *J. Neuropath. exp. Neurol.* **13**, 60–89 (1954).
- Waksman, B. H. & Adams, R. D. *J. exp. Med.* **102**, 213–236 (1955).
- Wisniewski, H., Prineas, J. & Raine, C. S. *Lab. Invest.* **21**, 105–118 (1969).
- Waksman, B. H. & Adams, R. D. *J. Neuropath. exp. Neurol.* **15**, 293–333 (1956).
- Raine, C. S., Wisniewski, H. & Prineas, J. *Lab. Invest.* **21**, 316–327 (1969).
- Madrid, R. E. & Wisniewski, H. M. *J. Neurocytol.* **7**, 265–282 (1978).
- Lassmann, H., Kitz, K. & Wisniewski, H. M. in *Search For the Cause of Multiple Sclerosis and Other Chronic Diseases of the Central Nervous System* (ed. Boese, A.) 96–104 (Verlag Chemie, Weinheim, 1980).
- Morrison, L. R. *Archs Neurol. Psychiat.*, Chicago **58**, 391–416 (1947).
- Rivers, T. M. & Schwentker, F. F. *J. exp. Med.* **61**, 689–702 (1935).
- Wisniewski, H. M. & Keith, A. B. *Ann. Neurol.* **1**, 144–148 (1977).
- Raine, C. S. in *Progress in Neuropathology* Vol. 3 (ed. Zimmerman, H. M.) 225–251 (Grune & Stratton, New York, 1976).
- Alvord, E. C. Jr, Shaw, C. M., Hruby, S. & Kies, M. W. *Ann. N.Y. Acad. Sci.* **122**, 333–345 (1965).
- Levine, S. & Sowers, R. in *The Suppression of Experimental Allergic Encephalomyelitis and Multiple Sclerosis* (eds Davison, A. N. & Cuzner, M. L.) 199–209 (Academic, London, 1980).
- Rasminsky, M. *Archs Neurol.* **28**, 287–292 (1973).
- Bostock, H., Sears, T. A. & Sherratt, R. M. *J. Physiol., Lond.* **313**, 301–315 (1981).

Differences in the kinetics of rod and cone synaptic transmission

Julie L. Schnapf* & David R. Copenhagen

Departments of Ophthalmology and Physiology,
University of California, San Francisco,
California 94143, USA

Photoreceptors of the vertebrate retina hyperpolarize in response to light. The hyperpolarization elicited by a brief flash is approximately ten times slower in rods than in cones of the same retina¹. We have examined the amplification and temporal properties of synaptic transfer of rod and cone signals to a common postsynaptic element, the horizontal cell. We find that the kinetics of signal transfer at these chemical synapses parallels the speed of the light-evoked signals themselves.

Synaptic transfer was studied by recording simultaneously the responses of pre- and postsynaptic cells to dim flashes of light. Eyecups from the snapping turtle (*Chelydra serpentina*) were prepared as described earlier². Two intracellular micro-electrodes, with tips ~100 µm apart in the plane of the retina, were independently lowered into the retina. Intracellular recordings were made from photoreceptors and from axon

terminals of luminosity horizontal cells (H1AT). H1ATs receive synaptic input from both rods and red-sensitive cones^{3,4}. Cells were classified on the basis of their receptive field properties and response waveforms, criteria which have been justified with intracellular dye injections^{2,4,5}. Changes in membrane potential were recorded in response to brief (13 ms) flashes of light at a wavelength of 520 or 650 nm. The light intensities used were shown experimentally to be in the linear range¹. Large circular spots of light (1.5–4.3 mm diameter) were utilized in order to eliminate temporal filtering of signals caused by coupling between photoreceptors and between horizontal cells^{6,7}. Six rod–horizontal cell pairs and five cone–horizontal cell pairs were used in this analysis.

Stimulation of cones with bright lights elicited membrane hyperpolarizations in the horizontal cells of up to 40 mV, while rods elicited horizontal cell hyperpolarizations of only a few millivolts. The cone- and rod-mediated inputs to horizontal cells were adequately separated by using stimuli of different wavelengths. The horizontal cell response to dim 520-nm test flashes was mediated solely by rods, as no measurable response in a cone was ever recorded to test flashes of the intensities used to stimulate rod–horizontal cell pairs. Although 650 nm light excited both rods and cones, the faster cone contribution to the horizontal cell response could be clearly separated from that of the rod⁴. To confirm that the initial part of the 650-nm

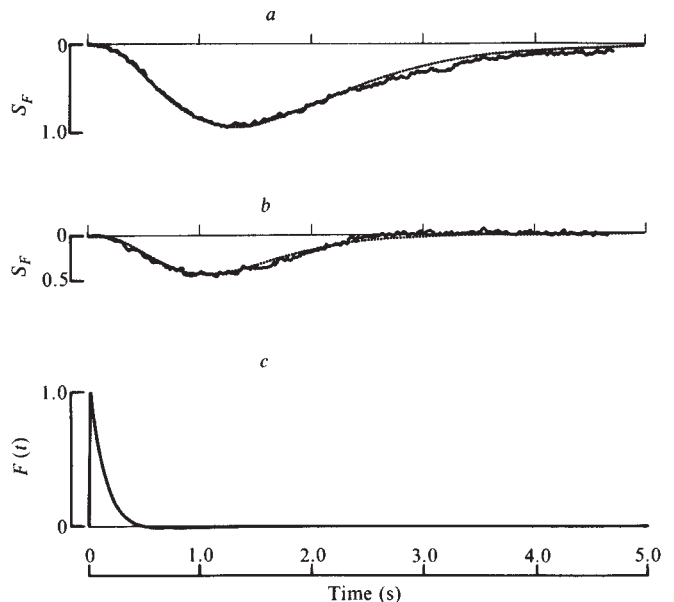


Fig. 1 The rod–horizontal cell synapse. Results from 63 test flashes, 520 nm, presented at $t = 0$. Two test flash intensities were chosen in order to test for intensity/response linearity. The intensity of the first 30 flashes was 1.0 photon per μm^2 per flash; the intensity of the second 33 was 0.5 photons per μm^2 per flash. Stimulus diameter 3.8 mm; S_F , in units of mV per photon per μm^2 per flash, was obtained by dividing the voltage responses by the flash intensity. *a*, Ensemble average of light responses of rod. Dashed curve fitted to average is the equation $S_F(t) = A(te/T)^n \exp(-tn/T)$, where A is the sensitivity at the peak of the light response in mV per photon per μm^2 per flash, T is the time to the peak of the response in seconds and n is an integer chosen to give the best fit by eye of the experimental data to the curve⁸. In some cases, a non-integer value of n gave a better fit to the rising phase of the light response, but the deconvolution procedure described below requires an integer value. However, the selection of an integer above or below this value has only a small effect on the duration of the calculated impulse response. $A = 0.90$, $T = 1.34$ and $n = 4$. *b*, Ensemble average of light responses of horizontal cell. Responses recorded simultaneously with rod responses for *a* above. Dashed curve fitted to average is the same equation as for *a*, with $A = 0.45$, $T = 1.06$ and $n = 5$. *c*, Normalized impulse response of rod–horizontal cell synapse. The curve is calculated by deconvolving the equations fitted to the horizontal cell and photoreceptor light responses. The solution to that deconvolution is

$$F(t) = C \exp(-tn_2/T_2) \sum_{k=0}^{n_1+1} t^{(n_2-k)} \frac{(n_1/T_1 - n_2/T_2)^{-k}}{(n_2-k)! k! (n_1+1-k)!}$$

where C is a constant chosen to normalize $F(t)$ to one at its peak, A , n and T are the constants described above, with subscripts 1 and 2 referring to the values for the photoreceptor and horizontal cell respectively.

* Present address: Department of Neurobiology, Sherman Fairchild Science Building, Stanford University School of Medicine, Stanford, California 94305, USA.

response was not contaminated by rod inputs, the 650-nm test flash was presented on dim red or green backgrounds whose intensities were adjusted for equal cone absorption. In these conditions, cone responses and the early portion of the horizontal cell responses were independent of the wavelength of the background light.

Averaged responses to dim test flashes are shown in Fig. 1 for a rod-horizontal cell pair, and in Fig. 2 for a cone-horizontal cell pair. In all pairs recorded, the response of the photoreceptor peaked after the response of the horizontal cell. On average, the peak of the cone response followed the peak of the cone-driven response of the horizontal cell by 8 ms (range 2–18); the peak of the rod response followed that of the rod-driven response of the horizontal cell by 425 ms (range 280–600). The gain of synaptic transfer (the ratio of peak horizontal cell amplitude to peak photoreceptor amplitude) was 5.0 for the horizontal cell-cone pairs (range 2–10) and 0.46 for the horizontal cell-rod pairs (range 0.4–1.2). These values are comparable with those obtained by sequential recording of cones and horizontal cells⁸, but smaller than the rod-horizontal cell gains previously described⁴.

To determine how the waveform of the light response of a photoreceptor is transformed during its transfer to the horizontal cell, we calculated the impulse response of the synapse, $F(t)$. This is the calculated form of the voltage response in the horizontal cell for a brief change in voltage in the photoreceptor, and it reflects the rate-limiting steps in synaptic transfer.

Figures 1c and 2c are plots of the normalized impulse response. The method for obtaining this function is described in Fig. 1 legend. The impulse responses of rod and cone synapses are similar in shape and differ only in time scale. The impulse response rises immediately at zero time and slowly relaxes back to baseline with a small undershoot. This waveform is characteristic of a leaky high-pass filter in series with a low-pass filter.

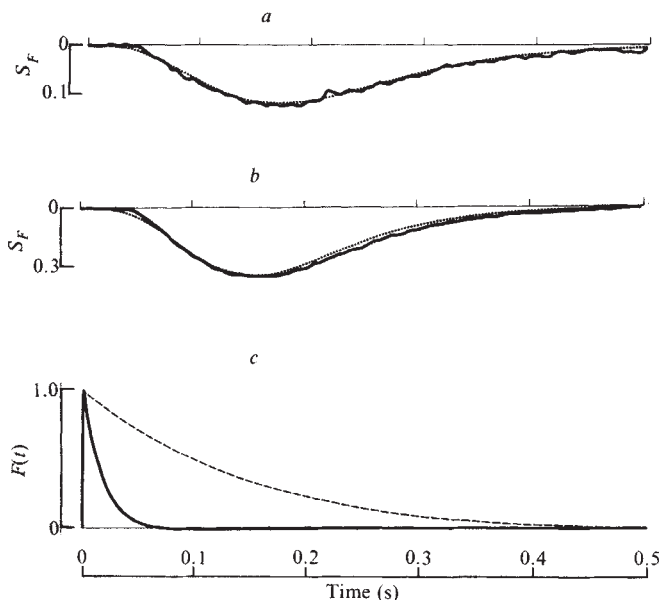


Fig. 2 The cone-horizontal cell synapse. Results from 53 test flashes, 650 nm, presented at $t=0$. The intensity of the first 23 flashes was 4.0 photons per μm^2 per flash, and of the second 30 flashes was 8.9 photons per μm^2 per flash. Stimulus diameter 3 mm, approximately twice the diameter of the receptive field of the horizontal cell. S_F is in units of mV per photon per μm^2 per flash. Note that the time scale has been expanded 10 times relative to Fig. 1. *a*, Ensemble average of light responses of cone. Dashed curve fitted to average is the equation in Fig. 1a legend with $A=0.115$, $T=0.174$ and $n=5$. *b*, Ensemble average of light response of horizontal cell. Responses were recorded simultaneously with the cone response for *a* above. Dashed curve fitted to average is the equation in Fig. 1a legend, with $A=0.345$, $T=0.156$ and $n=6$. *c*, Normalized impulse response of cone-horizontal cell synapse (the smooth curve). The curve is calculated from the equation in Fig. 1c legend using the values given in 2a and 2b as the constants of the equation. For comparison, the impulse response in Fig. 1 of the rod synapse is re-plotted on the faster time scale (---).

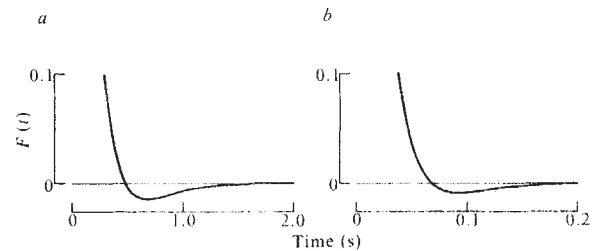


Fig. 3 Undershoot of the impulse response. Portions of the impulse response in Figs 1, 2 are redrawn at higher magnification to illustrate the small undershoot. *a* Is the impulse of the rod-horizontal cell synapse of Fig. 1, and *b* is that of the cone-horizontal cell synapse of Fig. 2.

High-pass filtering (or differentiation) is required to account for the faster time to peak of the horizontal cell response relative to the photoreceptor response; the high-pass filtering is also responsible for the negative component of the impulse response which can be seen to dip just below the baseline on the magnified scale in Fig. 3. Low-pass filtering is required to fit the rising phase of the horizontal cell response. At early times the light response of the photoreceptor rises as approximately t^n , while the horizontal cell response rises as t^{n+1} . (Each low-pass stage through which the photoreceptor signal is filtered adds an additional power of t to the rising portion of the horizontal cell response⁹.)

The nature of synaptic filtering is unclear. The high-pass filter stage may reflect time- and voltage-dependent behaviour of the postsynaptic channels¹⁰, postsynaptic 'desensitization'¹¹, or presynaptic depletion of transmitter¹² due to prolonged release. The low-pass filtering stage may represent the finite time that postsynaptic channels are opened by the transmitter released from photoreceptors¹³. In cone-horizontal cell pairs, the low-pass filtering might also simply reflect the filtering of the membrane voltage of the horizontal cell by its membrane time constant¹⁴. The kinetics of transmitter release, degradation and uptake at synapses in other physiological systems are fast relative to channel opening and membrane time constants^{13,15}, but these processes may be slower at the photoreceptor synapse.

The duration of the synaptic impulse is much greater for the synapse between rods and horizontal cells than between cones and horizontal cells: the impulse response falls to $1/e$ of its peak value in, on average, 131 ms for the rod synapse (range 100–160) and in 16 ms for the cone synapse (range 12–20). In comparison, in an analysis of the fluctuation of membrane potentials of bipolar cells in the same retina¹⁶, the impulse for the cone-bipolar cell synapse was estimated to last ~14 ms in the hyperpolarizing bipolar and ~50 ms in the depolarizing bipolar. A small degree of high-pass filtering was also observed in the bipolar synapses.

The time scale of the synaptic impulse from rod to horizontal cell is approximately 10 times slower than in the cone-horizontal cell pair. This difference parallels a difference in the speed of light responses of rods and cones. It was previously suggested that the match between the temporal properties of the light response of photoreceptors and the kinetics of synaptic transmission to ganglion cells might improve the signal-to-noise ratio of transmission in the retina¹⁷. The present experiments demonstrate that this kind of optimization is already apparent at the first synapses of the retina.

We thank Drs J. F. Ashmore, D. A. Baylor, B. J. Nunn and D. G. Green for comments on the manuscript, and Ms J. Ansello for typing the manuscript. This research was supported by NIH grants NS-07067 and EY-01869.

Received 19 November 1981; accepted 16 March 1982.

1. Baylor, D. A. & Hodgkin, A. L. *J. Physiol., Lond.* **234**, 163–198 (1973).
2. Copenhagen, D. R. & Owen, W. G. *J. Physiol., Lond.* **259**, 251–282 (1976).
3. Leeper, H. F. *J. comp. Neurol.* **182**, 795–810 (1978).
4. Leeper, H. F. & Copenhagen, D. R. *Vis. Res.* **19**, 407–412 (1979).
5. Baylor, D. A. & Fuortes, M. G. F. *J. Physiol., Lond.* **207**, 77–92 (1970).
6. Schwartz, E. A. *J. Physiol., Lond.* **257**, 379–406 (1976).

7. Detwiler, P. B., Hodgkin, A. L. & McNaughton, P. A. *Nature* **274**, 562-565 (1978).
8. Normann, R. A. & Perlman, I. *J. Physiol., Lond.* **286**, 509-524 (1979).
9. Fuortes, M. G. F. & Hodgkin, A. L. *J. Physiol., Lond.* **172**, 239-263 (1964).
10. Gage, P. W. & McBurney, R. N. *J. Physiol., Lond.* **244**, 385-407 (1975).
11. Katz, B. & Thesleff, S. *J. Physiol., Lond.* **138**, 63-80 (1957).
12. Furukawa, T. & Matsuura, S. *J. Physiol., Lond.* **276**, 193-209 (1978).
13. Anderson, C. R. & Stevens, C. F. *J. Physiol., Lond.* **235**, 655-691 (1973).
14. Johnston, D. & Lam, D. M. K. *Nature* **292**, 451-454 (1981).
15. Crawford, A. C. & McBurney, R. N. *J. Physiol., Lond.* **258**, 205-225 (1976).
16. Ashmore, J. F. & Copenhagen, D. R. *Nature* **288**, 84-86 (1980).
17. Baylor, D. A. & Fettiplace, R. *J. Physiol., Lond.* **271**, 425-448 (1977).

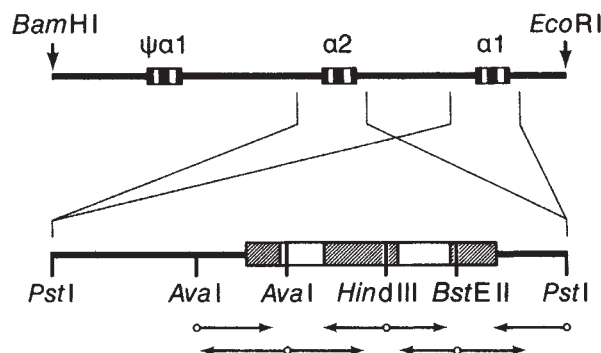


Fig. 1 Diagram showing the DNA segment cloned from the patient with non-deletion Hb-H disease, and the sequencing strategy of the α loci. Leukocyte DNA from the proband was digested with *Bam*HI and the 14.5-kb fragment isolated by discontinuous electrophoresis on an agarose gel¹⁴. The *Bam*HI end was made flush by adding the tetranucleotide GATC with Klenow fragments of DNA polymerase I, and was ligated to an *Eco*RI linker¹⁵. The DNA was then digested with *Eco*RI and the resultant 13-kb fragment containing the 5' $\psi\alpha$ - α 2 and α 1 3' loci was ligated to Charon 4A arms at the *Eco*RI site. The phage was packaged and used to infect *Escherichia coli* DP50 Sup F¹⁶. Phages containing α -globin inserts were identified by the method of Benton and Davis¹⁷ with ³²P-labelled α -globin cDNA, and those with a positive signal were screened twice more. One such isolate, λ GL6, was grown in a 1 l culture. The phage particle was isolated and banded on a caesium chloride gradient (1.5 g ml⁻¹) in a Ti50 rotor at 30,000 r.p.m. for 20 h. The DNA was then digested with *Eco*RI and *Bgl*II, and two fragments of 9.2 and 3.6 kb, which contained the $\psi\alpha$ - α 2- and α 1-globin loci respectively, were subcloned into the *Eco*RI-*Bam*HI sites of pBR322. The restriction sites used for sequencing are shown in the lower diagram. Sequencing was performed according to the Maxam and Gilbert method¹⁸, and carried out side by side with the corresponding fragments from the normal α -globin loci cloned by Lauer *et al.*¹⁹.

Globin structural mutant $\alpha^{125\text{Leu} \rightarrow \text{Pro}}$ is a novel cause of α -thalassaemia

Michel Goossens*, Kathleen Y. Lee,
Stephen A. Liebhaber* & Yuet Wai Kan†

The Howard Hughes Medical Institute Laboratory and the Division of Molecular Hematology, Department of Medicine, University of California, San Francisco, California 94143, USA

The duplicated human α -globin structural loci lie on chromosome 16 and are arranged in the order 5' α 2- α 1 3'. Although the most common molecular mechanism for α -thalassaemia is deletion of segments of DNA that contain one or both of these α -globin structural loci¹, two molecular lesions that are not due to gross gene deletion have been defined. In one, a point mutation in the termination codon of the α 2 gene results in the production of an elongated α -globin chain^{2,3}. As only a small amount of the mutant globin chain is produced, α -thalassaemia results. The second lesion, which is caused by a 5-base pair (bp) deletion in the first intervening sequence of the α 2 gene, results in abnormal mRNA processing and α -globin chain deficiency⁴. We now describe a novel mechanism for α -thalassaemia not involving deletion. A single nucleotide mutation in the coding region of the α 2 gene results in the substitution of proline (Pro) for leucine (Leu) in a region of the H helix of the α -globin chain, which is critical for α 1- β 1 contact. This probably impedes α 1- β 1 dimer formation, the initial step of haemoglobin tetramer assembly, and produces an α -thalassaemia phenotype.

The clinical, biochemical and genetic data on a particular patient with haemoglobin-H (Hb-H) disease and his family have been reported in detail elsewhere⁵. Usually only one of the four α -globin loci remains in Hb-H disease (genotype --/– α); the proband and his sister, however, have Hb-H disease with both α -globin loci deleted on one chromosome and intact on the other (genotype --/– $\alpha\alpha$). The family members who have microcytosis characteristic of α -thalassaemia have two different genotypes: the usual two-gene deletion (–/– $\alpha\alpha$) and an unusual genotype in which all four α -globin genes are intact ($\alpha\alpha/\alpha\alpha$). To explain the presence of α -thalassaemia trait in members of the latter genotype, we postulated that one of the chromosomes with two intact α -globin genes was defective in α -globin chain production [$\alpha\alpha/(\alpha\alpha)^T$], and that the combination of this non-deletion α -thalassaemia chromosome with the usual two-gene deletion form produced the Hb-H phenotype seen in the proband and his sister [–/– $(\alpha\alpha)^T$].

Previous Southern analysis of the proband's DNA revealed no abnormalities in the α -globin gene map⁶. mRNA analysis showed an α 1- to α 2-globin mRNA ratio of 2.5:1, which is identical to that found in normal reticulocytes, and indicates that both α -globin genes are active in transcription⁷. To define the molecular lesion producing non-deletion α -thalassaemia, we cloned the segment of his DNA that contained the two α -globin genes. This segment must be derived from the chromo-

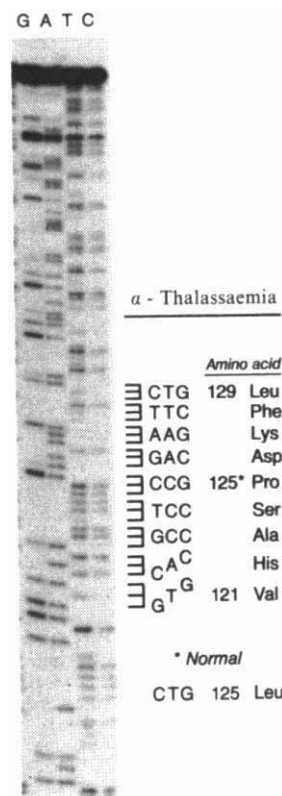


Fig. 2 Sequencing gel of the patient's DNA showing the region of the mutation. The DNA fragment was 5'-labelled at the *Bst*EII site with ³²P by polynucleotide kinase, and sequenced by the method of Maxam and Gilbert.

* Present addresses: Département de Biochimie, Hôpital Henri Mondor, Créteil, France (M.G.); Department of Human Genetics, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA (S.A.L.).

† To whom correspondence should be addressed.

some affected by non-deletion α -thalassaemia, as both α -globin loci are deleted from the other chromosome. We cloned in Charon 4A a 13-kilobase (kb) fragment extending from the *Bam*HI site to the *Eco*RI site and containing the $\psi\alpha$ -, $\alpha 2$ - and $\alpha 1$ -globin genes. The $\alpha 1$ - and $\alpha 2$ -globin loci were then separately subcloned into pBR322. Both the $\alpha 2$ - and the $\alpha 1$ -globin genes were sequenced side by side with the corresponding α -globin fragments from a previously isolated normal DNA. (See Fig. 1 for details.)

We completely sequenced the $\alpha 1$ and $\alpha 2$ loci, from the *Ava*I site located 100 nucleotides 5' to the cap site to the *Pst*I site 90 nucleotides 3' to the poly(A) addition site. The $\alpha 1$ -globin gene sequences from the propositus' DNA were identical with normal $\alpha 1$ sequences. We detected one difference, however, in the coding region of the $\alpha 2$ gene, at the position corresponding to amino acid 125, where the normal leucine codon (CTG) was changed to proline (CCG) (Fig. 2). We were able to confirm this mutation by restriction analysis because it abolished an *Eco*RI recognition site (CCTGG $\rightarrow$ CCGG) and created a new recognition site for the enzymes *Hpa*II and *Msp*I (CCGG). The normal 337-bp *Msp*I fragment that extends from the second intervening sequence to the 3'-noncoding region of the $\alpha 2$ -globin gene was cleaved into two fragments of 216 and 121 bp in the propositus' DNA. The mutation could potentially be detected in this patient's genomic DNA by methods similar to those used to detect the *Dde*I polymorphism in sickle cell anaemia⁸.

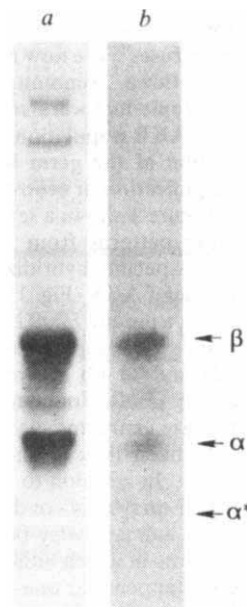
Leu $\rightarrow$ Pro mutations usually exert a profound effect on the stability of the haemoglobin tetramer. All the known Leu $\rightarrow$ Pro changes occurring in the α - and β -globin chains produce unstable haemoglobins and cause haemolytic anaemia⁹. For example, Hb Bibba, an α mutant also located in the H helix ($\alpha^{136}\text{Leu} \rightarrow \text{Pro}$), produces an unstable haemoglobin syndrome¹⁰. The unusual feature of the mutation in the family we studied is that the Leu $\rightarrow$ Pro mutation produces α -thalassaemia rather than an unstable haemoglobin. Carriers of this lesion (genotype $\alpha^{125}\text{Pro}/\alpha/\alpha$) are phenotypically indistinguishable from carriers of α -thalassaemia trait. The combination of this gene and a deletion α -globin gene ($--/\alpha^{125}\text{Pro}\alpha$) produces the clinical features of Hb-H disease.

A search for the abnormal α -globin chain using electrophoresis on starch gel, Triton-urea gel, and isoelectric focusing was unsuccessful. However, after incubating the patient's reticulocytes with ³H-leucine and separating the haemolysate on a Triton-urea gel, we observed an abnormal radioactive α -globin band that turned over rapidly (Fig. 3). This unstable α -globin could produce Hb-H disease by the following mechanism. Three amino acids ($\alpha^{122}\text{His}$, $\alpha^{123}\text{Ala}$ and $\alpha^{127}\text{Lys}$) in close proximity to $\alpha^{125}\text{Leu}$ participate in $\alpha 1$ - $\beta 1$ contact¹¹. The Leu $\rightarrow$ Pro substitution probably disrupts the H helix and in turn interferes with these contact points. The formation of $\alpha 1$ - $\beta 1$ dimers may thus be impeded and the ability of the β -chains to form Hb-H (β_4) may exceed that of the abnormal α -chain and the normal β -chain to form $\alpha 1$ - $\beta 1$ dimers. The uncombined α -chains would then be destroyed by proteolysis and an α -thalassaemia-like syndrome would ensue. Such a mechanism is consistent with the normal $\alpha 1$ to $\alpha 2$ ratio found in the propositus⁷ because the abnormal mechanism operates after translation. In contrast, in all cases of α - and β -thalassaemia produced by abnormal mRNA processing and by early or late termination, mRNA from the affected locus is either decreased or unstable *in vivo*^{4,7,12,13}.

In conclusion, we have described a human globin mutation, $\alpha^{125}\text{Leu} \rightarrow \text{Pro}$, that produces the α -thalassaemia phenotype through a novel post-translational mechanism. As it was the propositus' mother who carried the α -globin mutant, we have named this variant Quong Sze, after the province in China where she was born.

We thank Jennifer Gampell for editorial comments. This study was supported by grants from the NIH and the March of Dimes/Birth Defects Foundation. Y.W.K. is an investigator of the Howard Hughes Medical Institute.

Fig. 3 Electrophoresis of globin on a Triton-urea gel. The reticulocytes from the propositus with Hb-H disease were concentrated by Percoll-Reno M60 gradient²⁰ and incubated with ³H-leucine for 20 min as previously described²¹. Cells were lysed and 1 μ g of haemoglobin was applied to a Triton-urea gel and electrophoresed as described earlier²². The gel was stained with Coomassie blue, dried and autoradiographed for 4 weeks. Lane a, Coomassie blue staining of the gel; b, autoradiogram of the gel. The faint radioactive band (α^*) that migrates faster than the normal α -globin band in b has been characterized as the abnormal α -globin by hybrid arrest of translation by α -globin cDNA in a cell-free system (S.A.L. and Y.W.K., unpublished) and work is in progress to assess its synthesis and turnover rates.



Received 5 October 1981; accepted 8 March 1982.

- Weatherall, D. J. & Clegg, J. B. *The Thalassemia Syndromes* 3rd edn (Blackwell Scientific, Oxford, 1981).
- Milner, P. F., Clegg, J. B. & Weatherall, D. J. *Lancet* **i**, 729-732 (1971).
- Michelson, A. M. & Orkin, S. H. *Cell* **22**, 371-377 (1980).
- Orkin, S. H., Goff, S. C. & Hechtman, R. L. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5041-5045 (1981).
- Kan, Y. W., Dozy, A. M., Trecartin, R. & Todd, D. *New Engl. J. Med.* **297**, 1080-1084 (1977).
- Embury, S. H., Lebo, R. V., Dozy, A. M. & Kan, Y. W. *J. clin. Invest.* **63**, 1307-1310 (1979).
- Liebhafner, S. A. & Kan, Y. W. *J. clin. Invest.* **68**, 439-446 (1981).
- Chang, J. C. & Kan, Y. W. *Lancet* **ii**, 1127-1129 (1981).
- Bunn, H. F., Forget, B. G. & Ranney, H. *Human Hemoglobins* (Saunders, Philadelphia, 1977).
- Kleihauer, E. F. *et al. Biochim. biophys. Acta* **154**, 220-222 (1968).
- Vettore, L., DeMatteis, M. C. & Zampini, P. *Am. J. Hemat.* **8**, 291-297 (1980).
- Chang, J. C. & Kan, Y. W. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2886-2889 (1979).
- Trecartin, R. F. *et al. J. clin. Invest.* **68**, 1012-1017 (1981).
- Polsky, F., Edgell, M. H., Seidman, J. G. & Leder, P. *Analyt. Biochem.* **87**, 397-410 (1978).
- Goodman, H. M. *Meth. Enzym.* **65**, 63-64 (1980).
- Blattner, F. R. *et al. Science* **196**, 161-169 (1977).
- Benton, W. D. & Davis, R. W. *Science* **196**, 180-182 (1977).
- Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560-564 (1977).
- Lauer, J., Shen, C.-K. J. & Maniatis, T. *Cell* **20**, 119-130 (1980).
- Kan, Y. W., Schwartz, E. & Nathan, D. G. *J. clin. Invest.* **47**, 2515-2522 (1968).
- Alter, B. P. *Blood* **54**, 1158-1163 (1979).
- Lehmann, H. & Huntsman, R. G. *Man's Haemoglobins*, 449 (Lippincott, Philadelphia, 1974).

Germ-line MuLV reintegrations in AKR/J mice

Winship Herr & Walter Gilbert

Department of Biochemistry and Molecular Biology, Biological Laboratories, Harvard University, Cambridge, Massachusetts 02138, USA

AKR mice are viraemic from birth¹ and contract leukaemia at high incidence at 6-9 months of age². They express, in all tissues studied, the ecotropic (mouse infectious) retrovirus AKV¹. Rowe^{3,4} showed, by crossing AKR mice with the low-viraemia, non-leukaemic NIH/Swiss mice, that viraemia and the leukaemic phenotype are linked and carried at two independent loci (*Akv-1* and *Akv-2*). These two loci were shown to represent integrated ecotropic proviruses^{5,6}. As part of a study of the structure of proviruses which arise in the leukaemic thymus of AKR mice, we identified ecotropic-specific hybridization probes from different regions of the AKV genome⁷. Initial Southern hybridization⁸ analyses showed that individual leukaemic AKR/Jackson mice, which have been sibling inbred

for over 140 generations, had differing numbers of ecotropic proviruses⁹. We now report analyses showing that AKR/J mice had three common proviruses; however, two additional ecotropic loci were scattered throughout the Jackson Laboratory AKR population. These loci probably arose through reinfection of the germ line and represent a currently occurring amplification of ecotropic proviruses in these inbred mice.

Figure 1 shows a set of four different 'homozygous' hybridization patterns from individual AKR/J mice. The ecotropic virus-specific hybridization probe corresponds to the gp70 region of AKV (Fig. 1). As the AKV genome contains no *EcoRI* sites⁶, a unique *EcoRI* fragment, larger than 8.8 kilobases, arises from each integrated AKV provirus. AKR/J mice all carry 50-, 12- and 11-kb ecotropic *EcoRI* fragments (lanes a-d). The 50-kb *EcoRI* fragment is the *Akv-1* allele⁶, while the other two are characteristic of the AKR/J substrain⁶ and have probably been fixed within these AKR mice since the breeding began. In addition to these three loci we observe AKR/J mice which carry a 14- or 17-kb *EcoRI* fragment (Fig. 1, lanes b, c), some animals carry both proviruses (lane d). We also observe patterns in which either the 14- or 17-kb fragment is heterozygous (appears at one-half intensity, see Fig. 2 as an example). New ecotropic *EcoRI* fragments have also recently been observed in AKR/J mice by Yoshimura and Breda¹⁰.

All five proviral *EcoRI* restriction fragments seem to contain a full-length AKV or closely related genome. They are larger than 8.8 kb and each hybridizes to the three different ecotropic probes. (Figure 1 shows the three probes on a restriction map of AKV¹¹.) In addition, *PstI* digestion of each of the four different samples in Fig. 1 (*EcoRI* lanes a-d) produces a single full-length internal *PstI* fragment (Fig. 1, *PstI* lanes a-d), consistent with the restriction map of AKV.

Figure 2 shows that the 14-kb *EcoRI* ecotropic restriction fragment is inherited in the germ line. A female AKR/Jackson mouse homozygous at all five ecotropic proviral loci was mated with a male AKR/Jackson mouse which lacked the 14-kb proviral *EcoRI* fragment. *EcoRI* digestion and Southern blot analysis of the F₁ embryos revealed that the 14-kb fragment was present but at one-half the intensity of the female parent. An analogous experiment showed that the 17-kb fragment is also carried in the germ line (data not shown).

To examine the cellular sequences on each side of the provirus, we used a combination of enzymes, one of which cuts inside the provirus, and a specific probe that lies on a known side of the cut. Figure 3A shows, using the *pol* probe and a combination of *XbaI* and *EcoRI*, that the five 5' sequences are all different. Although the three fixed proviruses generate the same 13-kb *XbaI* fragment, this is fortuitous, because double digestion with *XbaI* and *EcoRI* generates three unique fragments (9.2, 10 and 13 kb), which shows that the *EcoRI* sites differ. The 17-kb *EcoRI* fragment generates unique 15-kb *XbaI* and 10.2-kb *EcoRI-XbaI* fragments whereas the 14-kb *EcoRI* locus produces an 11-kb *XbaI* fragment which contains no *EcoRI* sites. We also identified the 3' cellular *BamHI* and *SacI* sites with the C-p15E and gp70 probes, respectively. Figure 3B summarizes all the data as restriction site maps for the different proviral sites. As it is unlikely that these differences are due to point mutations, because we detect no alterations in the *PstI* restriction sites within the proviruses (see Fig. 1), we conclude that the cellular sequences around each provirus are different; these new proviruses probably arose through reinfection of germ-line cells and integration at unrelated sites.

We analysed 21 pedigreed mice from the AKR/Jackson pedigree expansion colony in January 1980, to examine the distribution of ecotropic proviruses. Figure 4 shows the relationship between the mice surveyed, outlined as a family tree, and the Southern blot analysis from which we determined whether the mice were heterozygous or homozygous at each locus. The variable bands are more prevalent in this survey than in an original pool of DNA from 20 mice examined in March 1978, when the variable proviruses were present in only 10-20% of the haploid genomes. The 17-kb fragment is present in 22 of

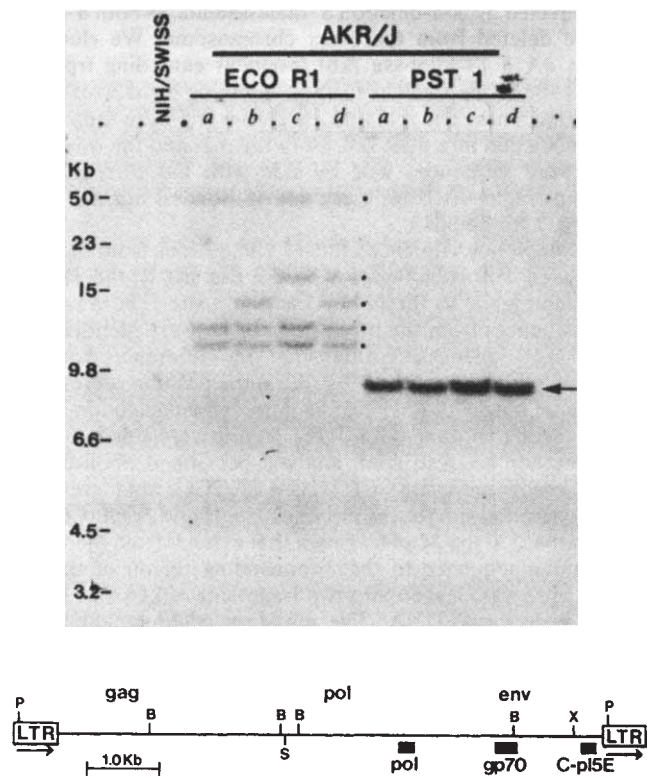


Fig. 1 Individual AKR/J mice, although sibling inbred for over 140 generations, do not carry identical numbers of integrated ecotropic provirus sequences. DNAs from four mice (a, b, c and d) previously identified as exhibiting different homozygous patterns were used in this Southern hybridization. Digestion with *EcoRI* generates the four different patterns whereas *PstI* digestion produces the single internal *PstI* fragment (arrow). The three constant *EcoRI* fragments are approximately 11, 12 and 50 kb long. The variable proviral *EcoRI* fragments are 14 kb (lane b) and 17 kb (lane c). Lane d displays all the ecotropic proviral fragments we have detected in AKR/J mice. 10 µg of purified¹⁵ brain DNA was digested with either *EcoRI* or *PstI*. The samples were electrophoresed through a 0.5% agarose gel¹⁶ using undigested λ phage DNA, λ phage DNA digested with *HindIII*, λ phage DNA digested with *XhoI*, and pBR322 digested with *BamHI* and *PstI* as size markers. The fractionated DNA was transferred to nitrocellulose filters^{8,17}, hybridized^{17,18} with the gp70 probe, washed¹⁸ and autoradiographed for 10 days with an intensifying screen^{19,20}. A control lane, labelled NIH/Swiss, contained 20 µg of *EcoRI* digested DNA isolated from the liver of an NIH/Swiss mouse, which does not contain a complete ecotropic provirus²¹ and shows, as expected, no hybridization. The location of the ecotropic-specific hybridization probes⁷ (black boxes) is shown with respect to a restriction map of the AKV genome¹¹: P, *PstI*; B, *BamHI*; S, *SacI*; X, *XbaI*. The AKV genome contains no *EcoRI* sites. The long terminal repeats (LTR) are shown as boxes. The probes are single-stranded radio-labelled AKV cDNA restriction fragments⁷ prepared by the method of D. Schwartz (personal communication). The p15E probe lies within the C-terminal coding region of p15E and is called C-p15E⁷.

the 42 haploid genomes examined and the 14-kb fragment is found in 19. These results are not an average of the entire mouse colony in January 1980, because the mice were selected such that they would be most distant from one another; nonetheless, it is likely that these new proviruses have increased in frequency between the two analyses. The pattern of provirus-containing loci displayed in Fig. 4 indicates that: (1) the common breeding pair at the top of the chart was heterozygous at both loci, (2) there is no obvious linkage relationship between loci and (3) both proviruses probably arose concurrently.

Different AKR substrains carry varying germ-line patterns of ecotropic proviruses in which only the *Akv-1* allele is common to all^{6,12}; this suggests that germ-line reintegrations have occurred in these mice and that the original AKR line contained only the *Akv-1* locus. Rowe and Kozak¹³ have observed that new virus-inducing loci appear in inbred *Akv-1* congenic mice if the maternal mice are viraemic. These are probably reintegrations of the AKV genome, which suggests that the new loci we

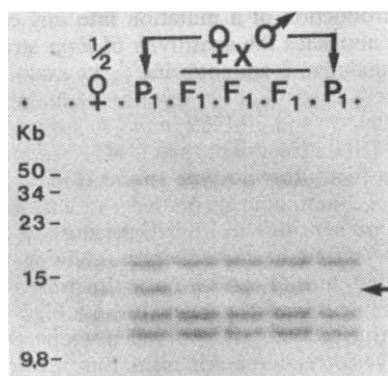


Fig. 2 Southern hybridization analysis of a cross between a male AKR/J mouse lacking the 14-kb proviral *EcoRI* fragment and a female AKR/J mouse homozygous for this fragment shows that these proviruses are carried in the germ line. The DNA from both parent AKR/J mice was isolated from brain tissue. The F_1 DNA was from 16-day old embryos. In each case, 20 μ g of *EcoRI*-digested DNA was loaded in each lane, except the left lane in which only 10 μ g of the female parent DNA was loaded. The filters were hybridized with the gp70 probe and autoradiographed for 10 days as described in Fig. 1 legend. The left two lanes contain the female parent DNA, the right lane contains the male parent DNA and the three middle lanes DNA from three individual embryos. The arrow indicates the 14-kb *EcoRI* fragment which is heterozygous in the embryo genome.

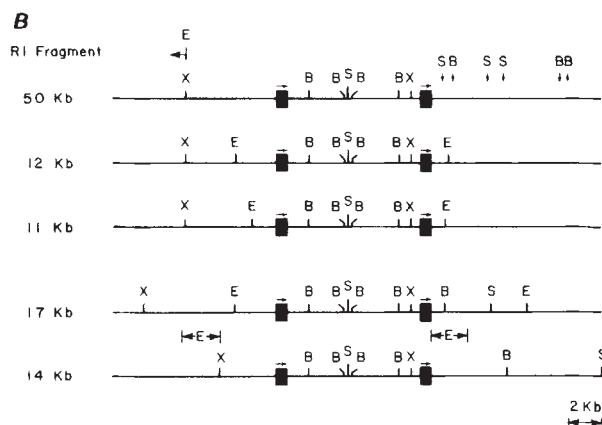
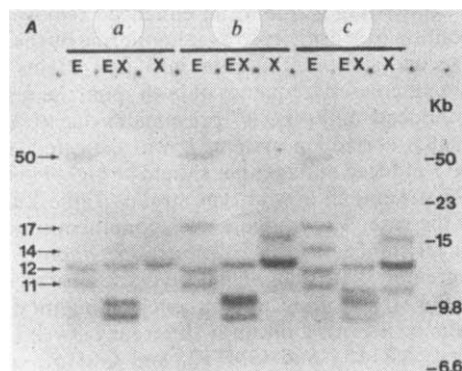


Fig. 3 Digestion of AKR/J DNA with *XbaI*, which cuts once within the AKV provirus, and *EcoRI*, shows that the cellular sequences surrounding the ecotropic germ-line proviruses differ. **A**, Three samples of DNA (10 μ g per lane) isolated from the brain tissue of individual AKR/J mice were cut with *EcoRI* (E) and *XbaI* (X) or with both enzymes (EX). Sample **a** contains only the three constant *EcoRI* fragments, sample **b** carries in addition the 17-kb *EcoRI* fragment, and sample **c** contains all five provirus loci. The filters were hybridized to the *pol* probe as described in Fig. 1 legend and autoradiographed for 3 days. **B**, Restriction site maps of the cellular sequences surrounding each of the germ-line AKV proviruses. The provirus sites are from Steffen *et al.*¹¹. The *XbaI* (X) and *EcoRI* (E) sites were determined from the data shown in **A**. The position of the 3' *SacI* (S) and *BamHI* (B) sites were determined from similar experiments using the gp70 and C-pl5E probes, respectively. We have not discriminated between the 3' cellular *BamHI* and *SacI* sites of the three constant proviruses.

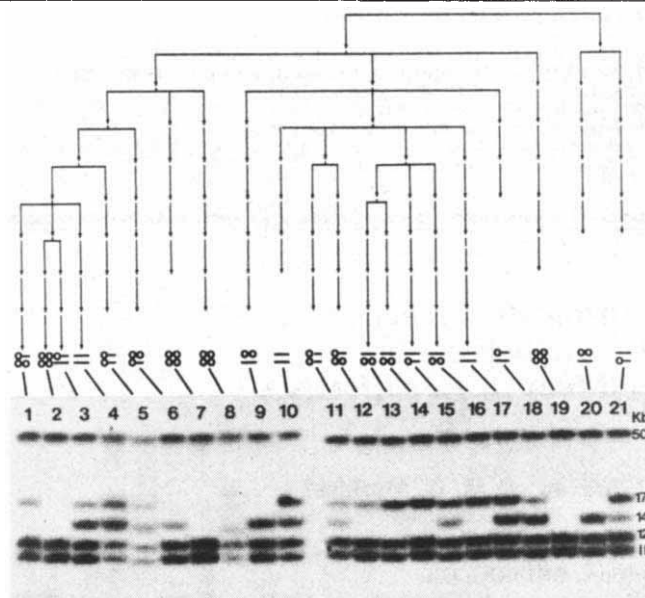


Fig. 4 Analysis of predigreed mice from the Jackson Laboratory colony reveals the distribution of the variable ecotropic proviruses. A pedigree chart was created for the pedigree expansion colony in January 1980, and 21 mice were selected to represent as wide a population as possible. The origin of each AKR/J mouse is shown in the upper portion of the figure. Each arrow indicates a generation. DNA (20 μ g) from each of 21 mice was cut with *EcoRI* and analysed with the gp70 probe as described in Fig. 1 legend. The interpretation of the autoradiogram is shown schematically above each lane. (0) Signifies the absence and (-) the presence of the variable provirus at a given locus. The filter was autoradiographed for 3 weeks. Samples 4, 5 and 8 contain partial digestion products.

observe in AKR/J mice may arise by reinfection of the embryo by AKV.

The AKR/J mice seem to be undergoing an amplification of the ecotropic proviruses. Xenotropic viruses might be the result of a past amplification of retroviral genomes. These viruses, which are unable to infect mouse cells but do replicate in cells from other species¹⁴, are present in much larger numbers (about 30 proviruses⁷) than the ecotropic proviruses. In a murine ancestor, the 'xenotropic' viruses may have replicated efficiently and amplified their numbers. By natural selection, the murine ancestor may have subsequently become resistant to infection by these viruses, creating a large number of relatively silent endogenous proviruses. Amplification of proviral genomes followed by mutation is one mechanism by which the organism could form a family of related genes, for example, the gp70 family coding for cell-surface markers.

We thank E. Les and Jackson Laboratories for help in screening the AKR/J mouse colony, D. Steffen for the NIH/Swiss DNA, M. Kreitman, N. Lonberg, D. Steffen and R. Weinberg for discussions, and D. Steffen for communicating results before publication. W.H. was supported, in part, by a NSF predoctoral fellowship. This work was supported by NIH grant GM09541-20.

Received 29 December 1981; accepted 22 February 1982.

- Rowe, W. P. & Pincus, T. *J. exp. Med.* **135**, 429-436 (1972).
- Gross, L. *Oncogenic Viruses*, 286-469 (Pergamon, Oxford, 1970).
- Rowe, W. P. *J. exp. Med.* **136**, 1272-1285 (1972).
- Rowe, W. P. *Harvey Lect. Ser.* **71**, 173-192 (1978).
- Chattopadhyay, S. K., Rowe, W. P., Teich, N. M. & Lowy, D. R. *Proc. natn. Acad. Sci. U.S.A.* **72**, 906-910 (1975).
- Steffen, D., Bird, S., Rowe, W. P. & Weinberg, R. A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4554-4558 (1979).
- Herr, W. & Gilbert, W. *J. molec. Biol.* (submitted).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
- Herr, W. & Gilbert, W. (in preparation).
- Yoshimura, F. K. & Breda, M. *J. Virol.* **39**, 808-815 (1981).
- Steffen, D. L., Bird, S. & Weinberg, R. A. *J. Virol.* **35**, 824-835 (1980).
- Quint, W., Quax, W., van der Putten, H. & Berns, A. *J. Virol.* **39**, 1-10 (1981).
- Rowe, W. P. & Kozak, C. A. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4871-4874 (1980).

14. Levy, J. A. *Science* **182**, 1151–1153 (1973).
15. Blin, N. & Stafford, D. W. *Nucleic Acids Res.* **3**, 2303–2308 (1976).
16. Helling, R. B., Goodman, H. M. & Boyer, H. W. *J. Virol.* **14**, 1235–1244 (1974).
17. Wahl, G. M., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683–3687 (1979).
18. Maniatis, T. *et al. Cell* **15**, 687–701 (1978).
19. Laskey, R. A. & Mills, A. D. *FEBS Lett.* **82**, 314–316 (1977).
20. Swanstrom, R. & Shank, P. R. *Analyt. Biochem.* **86**, 184–192 (1977).
21. Chattopadhyay, S. K., Lowy, D. R., Teich, N. M., Levine, A. S. & Rowe, W. P. *Proc. natn. Acad. Sci. U.S.A.* **71**, 167–171 (1974).

Mismatch correction at O⁶-methylguanine residues in *E. coli* DNA

P. Karran* & M. G. Marinus†

MRC Cell Mutation Unit, University of Sussex, Falmer, Brighton BN1 9QG, UK

† Department of Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts 01605, USA

Escherichia coli has a correction system which removes mismatched bases from DNA¹. Mutants (*dam*) which lack the major DNA adenine methylase² are hypersensitive to the effects of base analogue mutagens such as 2-aminopurine³ and appear to be defective in mismatch correction. Phenotypic revertants of *dam* mutants to base analogue resistance include second site mutations in *mutL* or *mutS* genes⁴, which are also part of the correction system. We report here that *E. coli dam* mutants are also sensitive to the DNA methylating agent *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG), which introduces O⁶-methylguanine (m⁶G) into the DNA. This sensitivity, however, was not observed using methylating agents which generate only low amounts of this alkylated base. Furthermore, the introduction of either a *mutL* or a *mutS* mutation into *dam* strains abolished the sensitivity to MNNG. These results suggest that mismatch correction occurs at m⁶G residues in DNA. These lesions miscode in DNA polymerase I-mediated DNA synthesis *in vitro*⁵ and are known to be mutagenic *in vivo*⁶. Nevertheless, it appears that mismatch correction at m⁶G residues in DNA does not lead to reduced induction of mutation by MNNG.

The major repair pathway for the promutagenic base m⁶G in *E. coli* is via a methyltransferase induced as part of the adaptive response to alkylating agents^{7–9}. The adaptive response is induced by exposure to sublethal concentrations of alkylating agents and renders cells resistant to the lethal and mutagenic effects of subsequent treatment with such agents⁹. To allow the maximum possible interaction between m⁶G and the mismatch correction system, except when specifically stated, all the experiments reported were performed in conditions in which the adaptive response is not induced. Figure 1a shows that the *dam-3* strain GM113 is more sensitive than its parental strain GM112 (*dam*⁺) to treatment with MNNG for 5 min. For MNNG, a fairly high proportion (7%) of the total alkylation products are m⁶G (ref. 10). To determine whether the sensitivity of *dam* strains to MNNG is related to the presence of this base, we tested the sensitivity of GM113 (*dam-3*) to dimethyl sulphate (DMS) for which m⁶G constitutes only a very minor proportion (0.5%) of the total products¹¹. Figure 1b shows that *dam* strains are not significantly more sensitive to DMS than *dam*⁺ strains; this agrees with a previous report that *dam* strains are only slightly hypersensitive to methylmethanesulphonate (MMS)¹² which also induces only a low proportion of m⁶G residues in DNA.

As the introduction of a mutation into any one of several mutator loci abolishes the sensitivity of *dam* strains to killing by the base analogue 2-aminopurine^{3,4}, we examined the effect on MNNG sensitivity of these second-site mutations. GM150 (*mutL-451 dam-3*) and GM169 (*mutS-453 dam-3*) were more resistant to MNNG treatment than GM113 (*dam-3*) and were similar in sensitivity to wild-type strains (Fig. 1c). By analogy to the effect of 2-aminopurine on these strains, we suggest that *dam* strains are sensitive to MNNG because of an attempted mismatch correction at base pairs containing m⁶G. In the absence of the requisite control normally present in wild-type (*dam*⁺) strains, this leads to cell death in a high proportion of cases. The introduction of the *mut* loci prevents the occurrence of this abortive correction mechanism, thus restoring wild-type survival levels. A reduction in the amount of m⁶G in *dam* strains should therefore result in a decreased frequency of these 'death-prone' repair events and thus enhanced survival after MNNG treatment. This reduction in the level of m⁶G may be brought about by pre-adaptation of the cells. Adapted *dam* strains should show enhanced survival after challenge with MNNG, but Jeggo *et al.*¹³ were unable to demonstrate such adaptation of *dam* strains to MNNG challenge. However, using a slightly different adaptation protocol, we have demonstrated enhanced survival in these strains after MNNG challenge (Fig. 2a). As normal levels of the m⁶G methyltransferase are induced in GM113 (*dam-3*) after adaptation (Fig. 2b), we conclude that at least part of the observed adaptation for survival in this strain is due to an enhanced removal of m⁶G.

The abolition of sensitivity to 2-aminopurine by the introduction of a second-site *mut* mutation into *dam* strains is accompanied by an increased frequency of both spontaneous and base analogue-induced mutagenesis¹⁴ presumably due to the absence of a mismatch correction system. It was expected, therefore, that MNNG-induced mutagenesis should be higher in the *mutL-451 dam-3* strain than in wild-type strains. Table 1 shows that this is not the case. The frequency of spontaneous mutation to rifampicin resistance of GM150 (*mutL-451 dam-3*) is ~100–200-fold greater than that of wild-type strains as expected. However, the MNNG-induced frequency of mutation to rifampicin resistance is not significantly different between the wild-type strains, AB1157 and GM112, and GM150 (*mutL-451 dam-3*). GM113 (*dam-3*) exhibits the smaller but consistent spontaneous mutator phenotype characteristic of *dam* mutants¹⁵ but again is not hypermutable by MNNG. It appears, therefore, that while the relative sensitivity to MNNG and DMS of GM113 (*dam-3*), GM150 (*mutL-451 dam-3*) and the parental GM112 (*dam*⁺) suggests that mismatch correction does occur in wild-type *E. coli* at m⁶G-containing base pairs, this correction system does not reduce miscoding mutagenesis due to this lesion. This interpretation is supported by the data of Sklar and Strauss¹⁶ which indicate that *E. coli uvrE* mutants, which are also deficient in mismatch correction, are normally mutable by MNNG.

Table 1 MNNG-induced mutation to rifampicin resistance

	Mutation frequency (Rif ^r mutants per 10 ⁸ bacteria)			
[MNNG] (μg ml ⁻¹)	0	5	10	20
Strain				
AB1157 (wild type)	2.6	2,900	4,900	33,000
GM112 (<i>dam</i> ⁺)	1.3	1,100	6,700	25,000
GM113 (<i>dam-3</i>)	94	1,100	3,900	9,600
GM150 (<i>mutL dam-3</i>)	270	2,000	6,200	23,000

Cultures (1 ml) containing ~10⁸ cells ml⁻¹ in supplemented minimal medium were treated with MNNG at the concentrations shown at 37 °C for 5 min. After treatment, cells were collected, washed and resuspended at 10⁷ cells ml⁻¹ in supplemented minimal medium and grown at 37 °C for 16–20 h. Appropriate dilutions were plated onto nutrient agar containing rifampicin (100 μg ml⁻¹). Total viable bacteria were determined by plating onto nutrient agar plates. The mutation frequency is expressed as Rif^r colonies per 10⁸ viable cells.

* Present address: Imperial Cancer Research Fund, Mill Hill Laboratories, Burtonhole Lane, Mill Hill, London NW7 1AD, UK.

Fig. 1 Sensitivity of *dam-3* strains to alkylating agents. Cultures in the exponential growth stage in minimal salts medium supplemented with 0.5% glucose, 0.1% casamino acids and 0.00003% thiamine were exposed to alkylating agent at the concentrations shown at 37 °C for 5 min. After dilution in 10 mM K phosphate pH 7.0, cells were plated onto nutrient agar. Survival was determined after growth at 37 °C for 24 h. *a*, MNNG sensitivity of ●, GM112 (*dam*⁺) and ▲, GM113 (*dam-3*). *b*, DMS sensitivity of ●, GM112 (*dam*⁺) and ▲, GM113 (*dam-3*). *c*, MNNG sensitivity of ■, GM150 (*mutL-451 dam-3*); ×, GM169 (*mutS-453 dam-3*); and ▲, GM113 (*dam-3*).

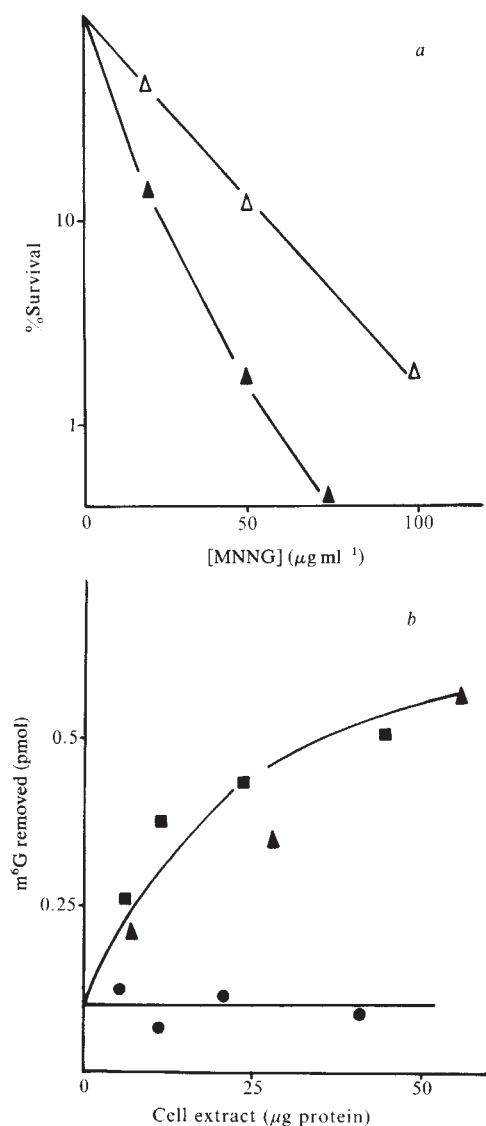
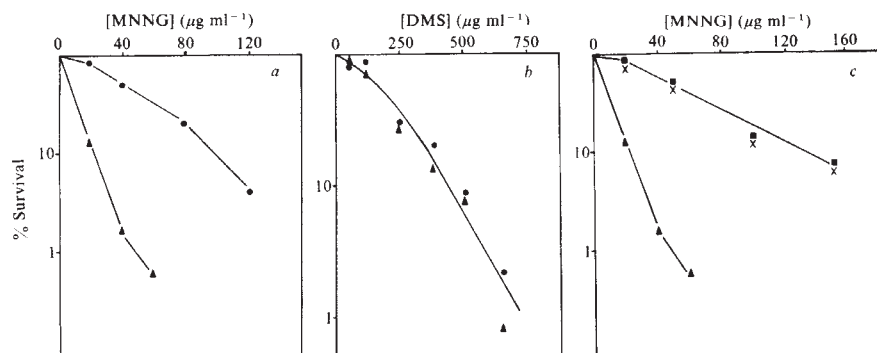


Fig. 2 Effect of adaptation on *dam* strains. *a*, MNNG sensitivity. Unadapted cultures (▲) were grown and treated as described in Fig. 1 legend. Adapted cultures (△) were grown in medium containing 0.1 $\mu\text{g ml}^{-1}$ MNNG for 3 h followed by 1 $\mu\text{g ml}^{-1}$ MNNG for 1 h immediately before treatment. *b*, Induction of *m*⁶G methyltransferase activity in adapted *dam-3* strains. Adapted and unadapted cultures were grown as described above and then cells were collected, washed and disrupted by sonication. The demethylation of ³H-labelled *m*⁶G in DNA treated with ³H-methylnitrosourea was assayed using these crude cell extracts as described elsewhere⁷ ●, GM113 (*dam-3*) unadapted; ▲, GM113 (*dam-3*) adapted; ■, GM1150 (*mutL dam-3*) adapted.

The ineffectiveness of the mismatch correction system in reducing MNNG-induced mutagenesis at *m*⁶G may be explained by the presence of the miscoding lesion in the parental DNA strand and by the ambiguous nature of *m*⁶G coding. After incorporation of 2-aminopurine or bromouracil, the 'incorrect' base remains in the daughter strand. In wild-type cells, the mismatch correction system will recognize the parental strand by its methylation pattern, and will thus always correct the mispair in the correct orientation. However, replication at *m*⁶G will direct the incorporation of T or C into daughter DNA. We propose that neither *m*⁶G:C nor *m*⁶G:T is a sufficiently good base pair to avoid subsequent recognition by the mismatch correction system. In normal conditions, the mismatch correction system will operate by removing the incorporated T or C, thus allowing a second attempt at the incorporation of the correct base. This second attempt will again result in an imperfect match as no 'correct' base exists. This removal/insertion cycle will presumably continue without check until the *dam*⁺ gene-encoded DNA methylase methylates the GATC sequences in daughter DNA thus removing the strand discrimination necessary for mismatch correction. If the incorporation of C or T opposite *m*⁶G occurs in approximately the same ratio for the polymerase system involved in replication and in mismatch correction, no net decrease in misincorporation and thus no decrease in mutagenesis will result from the operation of mismatch correction at *m*⁶G-containing base pairs.

There are, however, circumstances in which operation of mismatch correction at base pairs containing *m*⁶G should be effective in avoiding mutation. If the *m*⁶G methyltransferase operates during base removal and replacement and before the action of the *dam* encoded DNA methylase, the resulting G:C or G:T base pair will be either recognizably correct or a true mismatch and dealt with accordingly. In this case, the adaptive response, in addition to its role in removing *m*⁶G residues before DNA replication, may interact with the mismatch correction system to further reduce mutation frequency.

We thank Dr Michael Green for stimulating discussion throughout this work, and Professor B. A. Bridges and Dr A. R. Lehmann for critical reading of the manuscript. M.G.M. was the recipient of a Faculty Research Award (FRA-149) from the American Cancer Society.

Received 21 September 1981; accepted 25 February 1982.

1. Nevers, P. & Spatz, H. C. *Molec. gen. Genet.* **139**, 233-243 (1975).
2. Marinus, M. G. & Morris, N. R. *J. Bact.* **114**, 1143-1150 (1973).
3. Glickman, B., van den Elsen, P. & Radman, M. *Molec. gen. Genet.* **163**, 307-312 (1978).
4. Glickman, B. & Radman, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1063-1067 (1980).
5. Abbott, P. J. & Saffhill, R. *Biochim. biophys. Acta* **562**, 51-61 (1979).
6. Loveless, A. *Nature* **223**, 206-208 (1969).
7. Olsson, M. & Lindahl, T. *J. biol. Chem.* **255**, 10569-10571 (1980).
8. Schendel, P. F. & Robins, P. E. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6017-6020 (1978).
9. Samson, L. & Cairns, J. *Nature* **267**, 281-283 (1977).
10. Lawley, P. D. & Thatcher, C. J. *Biochem. J.* **116**, 693-707 (1970).
11. Harris, G., Lawley, P. D. & Olsen, I. *Carcinogenesis* **2**, 403-411 (1981).
12. Marinus, M. G. & Morris, N. R. *Mutat. Res.* **28**, 15-26 (1975).
13. Jeggo, P., Defais, M., Samson, L. & Schendel, P. *Molec. gen. Genet.* **162**, 299-305 (1978).
14. Rydberg, B. *Mutat. Res.* **52**, 11-24 (1978).
15. Bale, A., d'Alarcao, M. & Marinus, M. G. *Mutat. Res.* **59**, 157-165 (1979).
16. Sklar, R. & Strauss, B. *J. molec. Biol.* **143**, 343-362 (1980).

MATTERS ARISING

Early hominids and fire at Chesowanja, Kenya

IN addition to providing valuable new information on the important palaeo-anthropological evidence from Chesowanja, Gowlett *et al.*¹ state that "burnt clay found at one artefact locality dated to greater than 1.42 ± 0.07 Myr is the earliest known evidence of fire associated with a hominid occupation site" and that "the new find, along with the more tentative evidence from other sites, greatly strengthens the hypothesis that by 1.4 Myr hominids were using and controlling fire".

There are various reasons why the evidence reported should be treated as a good deal less definite than the article implies.

Traces of bush fires of presumably natural origin are not uncommon in early and middle Pleistocene deposits of the East African Rift Valley. For example, the fluvial facies of the Upper Member at Koobi Fora contains numerous reddened, hardened patches that are closely analogous to places where smouldering logs and stumps from modern bush fires have burned down against or into contemporary soils. I have seen and photographed several recent examples of burned patches in East Africa, and at least one rather analogous set of circumstances has been reported in Australia². Where these reddened patches, ancient or modern, erode, a scattering of hard red fragments results. Clasts of this natural 'terracotta' can commonly be observed in the grits and gravels formed on ancient land surfaces. I have frequently observed these clasts in excavations both at Koobi Fora and at Ologesailie and imagine them to be a widespread phenomenon.

This means that before a new record for a high antiquity of human control over fire can fairly be claimed, there must exist some objective means of distinguishing a hearth controlled by hominids from the baking effects of a bush fire. Gowlett *et al.* seem to claim that this distinction is made possible by the contrast between a determination of a 400 °C baking temperature for the burned material from site GnJi 1/6E compared with over 600 °C for "baking around a recently burned tree stump". I am sceptical that a difference between two isolated determinations can be relied on as a general criterion, and I would predict that if a series of bushfire baked samples is measured, the range of temperatures will be found to overlap with campfire temperatures.

The most convincing evidence for human control over fire in an open site would come from evidence of a localized, burned patch that was not a burned stump, inside the confines of artefactual refuse which had not been unduly distur-

bed by fluvial transport³. However, paragraph 4 and Fig. 2 of the article¹ make it clear that all the finds, including the burned earth fragments, were recovered from material that was swept together in the bed of a small channel. The finds can thus be regarded as provocative and suggestive, but no more.

The question of the antiquity of human control over fire is of more than curiosity value in our understanding of prehistory and human evolution. Besides the indication provided of an advance in mental capacity, fire may well have had important effects on feeding strategies and diet breadth. Control over fire also would have allowed humans to raise the frequency of bush fires and thereby to have a marked effect on vegetation patterns.

The article by Gowlett *et al.*, and others that they cite, serve the useful purpose of emphasizing that we still do not know whether humans controlled fire before ~0.5 Myr and if so, how long before. I suspect that development of suitable discriminants between traces of bush fires and controlled fire will require collaboration of physicists and archaeologists in the field during excavations.

GLYNN ISAAC

*Department of Anthropology,
University of California,
Berkeley, California 94720, USA*

1. Gowlett, J. A. J., Harris, J. W. K., Walton, D. & Wood, B. A. *Nature* **294**, 125–129 (1981).
2. Chapman, V. *Mankind* **11**, 480–483 (1977).
3. Laloy, J. *Cahiers du Centre de Recherches Préhistoriques* (Université de Paris I, 1980).

GOWLETT, HARRIS AND WOOD
REPLY—Isaac provides us with an opportunity to explain further why we consider that the context of the burnt clay at site GnJi 1/6E deserves serious consideration as evidence linking fire with hominid activity.

Isaac describes a series of circumstances which would provide the least ambiguous evidence of controlled fire on a single open site, and few would dispute his general criteria. We agree, too, that circumstances where baking is not proved, or where evidence of hominid activity is tenuous, provide inadequate testimony. Where we clearly disagree with Isaac is over the idea that a site with much stronger evidence should have no effect on what can be admitted as a working hypothesis.

We emphasize again that the evidence of fire at Chesowanja is quite definite: the clay was burned, and its association with the artefacts is direct and physical. Isaac rightly points out that the apparently low baking temperature of the clay is inconclusive evidence of the nature of the fire. This we clearly acknowledged, and it is why we advocated studies of the cooling

rate of the clay. We share with Isaac the hope that future investigations of experimental fires will assist in the interpretation of such data.

Strength of association is an important point. Given the controversy surrounding the contexts of all early sites¹, Isaac legitimately raises the possibility that the baked clay and artefacts may have been swept together. The evidence does not, however, support this view. On the contrary, we have good reason to believe that a reverse process has operated, and that the clay lumps, starting together, have crept apart. The presence of several large lumps together (square 82-83 E/111-112 S), including three weighing 262, 217 and 175 g respectively and which retain sharp edges and protuberances, argues very strongly against any significant water transport, as indeed does the whole context of fine-grained sediments. These finds occur close to a modern erosion gully, which prevents further exploration to the north and west, so that the scatter of smaller clay lumps to the opposite corner of the excavation is actually a vital bonus to interpretation, as it rules out any possibility of the burnt material deriving from recent superficial disturbance.

We welcome Isaac's contribution, for it is important that the cases for and against early human control of fire, at Chesowanja or elsewhere, should be fully aired. We also agree that progress in this field of research will be much facilitated by closer practical collaboration between physicists and archaeologists. Nonetheless, we stand firmly by the view that the evidence at Chesowanja greatly strengthens the hypothesis that fire was associated with hominid activities more than 1 Myr ago. We shall present additional evidence in due course.

We hope that this exchange will help to stimulate research on a wider scale, for although the evolutionary importance of controlled fire has been appreciated for over a century², new technical means for exploring its history now provide increasingly better chances of adding to the facts.

J. A. J. GOWLETT

*Research Laboratory for Archaeology
and the History of Art,
Oxford University, Oxford OX1 3QJ, UK*

J. W. K. HARRIS

*Department of Anthropology,
University of Wisconsin,
Milwaukee, Wisconsin 53201, USA*

B. A. WOOD

*Department of Anatomy,
The Middlesex Hospital Medical School,
London W1P 6DP, UK*

1. Binford, L. R. *Bones: Ancient Men and Modern Myths* (Academic, New York, 1981).
2. Darwin, C. *The Descent of Man* (Murray, London, 1871).

BOOK REVIEWS

Rolled over by technology?

Bruce Williams

THE God that limped? Hephaestus, the Greek god of fire and metalworking, was eagerly sought after for his technological skills. But he was ugly, irascible, often disliked by the other gods, and he had a pronounced limp. For Colin Norman this makes Hephaestus the appropriate symbol for modern technology — powerful and versatile but often marred by crippling defects.

A.N. Whitehead once wrote that the greatest invention of the nineteenth century was the invention of the art of invention. During the Second World War, J.A. Schumpeter predicted that industrial innovation would become a routine function of organized R & D. Planned inventions of weapons during the War, and a great range of planned inventions after it — sputnik, rockets to put men on the Moon and bring them back again, computers, chips, a great range of pesticides and ethical drugs, high-yield grains to create the green revolution — encouraged many to believe that the twentieth century had invented the science of invention. But that euphoria was undermined by the extent of serious problems created or intensified by the massive increases in R & D activities (the unforeseen side effects of new drugs, radioactive wastes, the displacement of labour), by fears that some of the solutions are clever but ephemeral (boosting crop yields from oil-based fertilizers), and by growing doubts about the capacity of R & D to contribute much to urban decay, poverty in the Third World, de-skilling, technological unemployment and the exhaustion of non-renewable resources.

While recognizing the immense social, economic and cultural benefits that science and technology have brought to society, Norman concentrates more on the social and environmental costs in the global economy. This is because he judged that present policies and trends are dangerous — the great post-War boom was built on an unstable foundation of cheap oil, and even now R & D in North America is more suited to the military needs of the 1950s than to the social needs of the 1980s. Whether the mismatch between defence and social needs in R & D objectives is any greater now than in the 1950s is not explored but is implied in his comments on the energy crisis and the problems of producing sufficient food, shelter and material resources for a population that will swell to six billion by the year 2000.

For those directly involved, there has

The God That Limp: Science and Technology in the Eighties. By Colin Norman. Pp.224. ISBN 0-393-01504-1. (W.W. Norton: 1982.) £10.50, \$18.25.

always been an element of fear associated with the introduction of new technologies. Norman refers to the growth of a more general fear that social changes are determined by technological change that is driven chiefly by its own momentum — “we even speak of new generations of computers, automobiles and other high technology goods as if they were biological descendants of earlier models”. He explains these growing fears in terms of the complexity of many modern technologies and the centralization of decision-making by both Governments and large corporations. On “the fear that society is simply a product of its technology”, he points to the major part that Governments play in determining the extent and directions of R & D and the numbers of scientists and engineers, and to the vast amount of innovation in, for example, weapons, space vehicles, law enforcement technologies, health care systems, as well as to the research activities in astronomy, high-energy physics and molecular biology, that cannot be explained even in terms of economic forces. Technological, economic and political forces interact to produce a complex array of pressures that push and pull technological developments along certain paths. The technical imperatives only define what is possible.

How, then, may technological development be pushed and pulled on to the path leading to a humane and sustainable society? He suggests four things — first, a better balance between unfettered market forces and Government constraints on market forces to reduce social costs; second, a broader public participation in the decisions that lead to the generation and adoption of new technology to give those most likely to be affected a role in its planning and to reduce some of the negative impacts; third, a greater access to modern technology in developing countries, and on more favourable terms, and an increase in their capacity to generate and apply technologies most suited to their own needs and resources; and fourth, a recognition that technology cannot by itself solve social and political problems.

His first suggestion is based on the judgement that the unfettered workings of the market system are enormously powerful in stimulating the development

and application of some technologies but cannot always be relied on to steer technological developments along socially appropriate paths. There is therefore a need for Governments to establish certain constraints — on pollution and working conditions — and to supplement market forces by subsidies for the production of certain commodities or services (including R & D). Government constraints on and supplements to market forces have been established for many years, the fact that competition is a form of social control based on legislation has been firmly accepted, and the theory of how to use taxes and bounties to control market forces when private and public costs or benefits diverge was established by Pigou 50 years ago. That Governments have not taken the appropriate actions — and Norman recognizes that some of their regulatory actions (on, for example, oil and gas prices in the USA) have made things worse — is sometimes due to the power of special interests, sometimes to unwisdom, and sometimes to the inability of scientists or engineers to provide firm evidence on the costs and benefits of various proposals for regulatory action. Norman writes of the urgent need for regulatory reform but his account of just what should be done and how is too cursory.

Norman admits that the experience of public participation in decision-making on new technologies is limited, but he is optimistic that it would have a beneficial effect on technological development and reduce some of the negative impacts.

He gives some examples of effective public participation in decisions in the USA, for example the Alaska Pipeline and the SST, and it is reasonable to demand and expect improved environmental impact statements and public discussions of them. But he underestimates the problems of participation. With new printing methods, for example, at what stages in the development of several technologies could there have been “public participation”? And when the publishers decide to use the labour-saving new technology who is to be “the public” to participate in the decisions? The new technologies that affect an industry may be developed and established outside that industry or in another country. The “compulsions of competition” are very complex, the effects are frequently impossible to foresee and there is need for much more consideration of how to distribute the costs of technical change.

His suggestions for providing the Third World with technology on better terms, for cooperation with the developing countries in the transfer and adaptation of technologies, and for the indigenous development of appropriate technologies are more positive and complete.

The God That Limps is a timely and stimulating book. In warning of the need to recognize that technology cannot of itself

solve social and political problems Colin Norman allows a much greater role for technology than Jay Forrester, who recently asked whether for the past 200 years technology has done much more than cope with the consequences of population growth. □

Sir Bruce Williams is Director of The Technical Change Centre, London, and Visiting Professor at Imperial College, University of London.

Meteorology enters the mesoscale era

Edward J. Zipser

Meso-scale Atmospheric Circulations. By B. W. Atkinson. Pp.495. ISBN 0-12-065960-3. (Academic: 1981.) £32.40, \$78.

FIFTY years hence, it may be difficult to explain to students how most meteorologists contrived to postpone to the 1980s their discovery of the importance of the mesoscale, the intermediate scales of motion. Indeed, how could so many observers of the natural world have accepted for so long that local weather events were controlled by large-scale systems with a spectral gap in the mesoscale (tens to hundreds of kilometres)? While better observations and numerical models have improved forecasts on time scales of 1–3 days, our ability to forecast 1–12 hours into the future, or to anticipate severe weather or flooding rains, has stagnated for 20 years.

Atkinson has written the first book in mesoscale meteorology, and his pioneering effort sets a high standard. He has made eminently reasonable selections of what to include and what to exclude. Each subject is treated in a similar manner, with observational knowledge summarized first, followed by theories, numerical simulation results and, where applicable, hardware models.

Many of the subjects have never been summarized comprehensively before. The first half of the book is a superb, systematic review of topographically induced flows, including lee waves, downslope winds, sea- and land-breeze circulations, and slope and valley wind circulations. The literature has been surveyed exhaustively; numerous papers hitherto unknown to this reviewer are placed in context and, where appropriate, recent papers highlighted. It will be difficult to improve upon this treatment, and Atkinson's book could well become the classic reference work on these subjects for the next 20 years.

Severe local storms have been the centrepiece of mesoscale meteorology for many years. Storms which generate hail, tornadoes and other damaging winds are challenging subjects because they are intrinsically scale-interaction phenomena. The severe storms chapter is excellent on convective storms in general, and in

summarizing recent literature on hailstorms, but it is rather weak on the rotational aspects and on recent literature on tornadic storms. The recent work on squall lines is well-covered, but it should have been clearly separated from the severe storms chapter and placed in a section on mesoscale convective systems, which are larger than individual storms. By not distinguishing between *organized* convection and *severe* convection, as where squall lines are claimed to be severe by definition, Atkinson's book will help some long-held misconceptions to persist for a few more years.

After a useful chapter largely concerned with recent results on shallow cellular circulations, the book closes by treating mesoscale circulations within cyclones. Atkinson himself has made contributions to extratropical cyclone studies, and this section is strong. Because this type of research concentrates upon radar description of precipitation distribution, however, it may not be obvious to all readers when it is proper to make inferences about mesoscale wind circulations in the absence of direct measurements. In addition, the section on tropical cyclones, like the literature on which it is based, is somewhat uneven. The recent improvement in aircraft instrumentation should lead to the need to update this section in the near future. For unknown reasons, no mention is made of cyclonic circulations in the tropics other than tropical storms.

This book does not attempt to be all things to all people; it deals with basic science, not applied science. The sea breeze, for example, is treated as a basic fluid flow, and the importance of that flow for precipitation distribution and air quality, although mentioned, is covered either briefly or not at all.

The author can take pride in fulfilling a long-standing need for a text in mesoscale meteorology, which is sure to stand as a basic reference work for decades. Nearly all meteorologists will want a copy close at hand. □

Edward J. Zipser is a Senior Scientist in the Convective Storms Division at the National Center for Atmospheric Research, Boulder, Colorado. (NCAR is sponsored by the National Science Foundation.)

People and change

William Brass

Population and Technology. By Ester Boserup. Pp.225. UK ISBN 0-631-12817-4 (Basil Blackwell: 1981); US ISBN 0-226-06673-8 (University of Chicago Press: 1981.) £12.50, \$17.50.

IN the early 1960s Ester Boserup began to propagate the view that population growth was one of the fundamental causes of the development of new agricultural methods. The importance of her work was less in the idea than in the lucidity and coherence of the demonstration. The present book is a wider essay on a similar theme, the interaction of population growth and technological change.

The central idea is that population factors have been a major determinant of the introduction of (or failure to exploit) new methods for the organization and improvement of production. Since the scope of the book is the whole of human history, much of it is logically concerned with agricultural output. The presentation is not of a highly structured argument from the idea to the consequences; rather there is an extensive historical review beginning with the transition from hunting-gathering to agriculture, and ending with the population-technological relationships in the less-developed countries of the present day. In between there are accounts of the establishment of urban civilizations, with particularly interesting sections on Mesopotamia, the Maya and China, the development of trade and industry in Europe, mass-migration to the United States and elsewhere, the transfer of industrial technologies to Asia and much else. Despite the unifying theme there is no dogmatism; population growth may be taken as the driving force but the importance of effects in the reverse direction is implicitly accepted.

Again and again in traditional works of pre-history and history there is a sentence which states that as a consequence of technological change (settled agriculture, spread of the use of iron, industrial invention) the population grew. There is seldom even a cursory consideration of how this happened; the way in which the changes could have altered fertility and/or mortality appropriately is left obscure. The large increase in the knowledge of the basic demography of less-developed communities in recent years has intensified the difficulty of the explanation. No simple, universal link between technological development and reduced mortality or increased fertility can be demonstrated.

The God that Limps: Science and Technology in the Eighties, by Colin Norman (reviewed on page 871) will be published in Great Britain on 26 May 1982.

The turn-around in the direction of the pressures, proposed by Ester Boserup, is stimulating and attractive although it leaves the causes of population growth in darkness, almost as a natural law of human history.

In a relatively short book of such a wide scope the evidence to support the thesis is inevitably condensed and the statements sweeping. Readers with a specialist knowledge of particular topics are likely to find assessments which do not take account of the most up-to-date research; the publications relied upon are not always the most reputable. For example, the discussions of the basic demography of the biosocial determinants of fertility and mortality, which are of critical significance for the reversal of the older views on the direction

of the relation between population growth and technological change, are surprisingly weak in depth of knowledge and citations of relevant research. Nevertheless they are sensible and only marginally misleading.

This is an exciting book, full of ideas which are provocative because they are unconventional and yet plausible. One of many possible examples is the examination of the minimum population densities needed for the development of urbanization at the relevant stage of technology. There is undoubtedly much to criticize in detail, but only by studies of this kind can the essential broad syntheses be achieved. □

William Brass is Professor of Medical Demography and Director of the Centre for Population Studies at the London School of Hygiene and Tropical Medicine.

The doctor of revolution as letter writer

Gavin Bridson

The Letters of Erasmus Darwin. Edited by Desmond King-Hele. Pp.363. ISBN 0-521-23706-8. (Cambridge University Press: 1981.) £45, \$95.

THAT remarkable polymath, Dr Darwin, has been singularly fortunate in having such a loyal admirer in Desmond King-Hele who has stayed beside his hero for some 20 years. His introduction to *Erasmus Darwin* (1963), followed by *The Essential Writings* (1968), *Doctor of Revolution* (1977), introductions to reprints of Darwin's *Botanic Garden* and *The Temple of Nature* (1973), and articles in *Nature* and elsewhere, all demonstrate this loyalty. His high regard for the amiable, versatile and near perfect eighteenth-century genius is no secret. In *Doctor of Revolution* he declared

Though I may be biased, I regard Erasmus Darwin as the greatest Englishman of the eighteenth century. If you disagree, can you name anyone else in the past 250 years with a list of accomplishments so numerous, so notable and so varied?

These he summarized in a list of 75 achievements which included such unsuspected topics as abolition of slavery, air travel, biological pest control, copying machines, oil drilling, rocket motors, speaking-machines, submarines, water closets and women's lib. He optimistically added that

The flow of specialised articles in journals has been increasing, and, if the momentum continues, we may yet see Darwin receiving the kind of scholarly attention that has been lavished on lesser figures like Boswell or Horace Walpole.

A glance at the *Isis* bibliographies shows that the flow has not been so great and that King-Hele's writings still provide a major

tributary. Little surprise, then, that he himself should edit Darwin's letters.

After a wide search for manuscripts in Britain and the United States, King-Hele believes that he has located the great majority of surviving letters and that this "provides a balanced and reasonably complete picture of Darwin the letter-writer in all his moods". For a man whose published writings ran to a million and a half words it is disappointing to find that the picture is composed of only 272 examples written between 1749 and 1802. No guess is made as to the likely quantity that may once have existed. We find that Boulton, Watt and Wedgwood kept 103 letters between them, and his grandson, Charles, quoted from 30 out of a collection that has not survived in manuscript. We must be grateful that as many as 88 other recipients make up the rest, that they spread over all but a handful of Darwin's 50 or so adult years, and that they reflect so much of the man and his achievements — for about half of King-Hele's list of 75 are apparently touched upon here. 187 letters survive in manuscript (162 in Darwin's own hand) and the rest are taken from published sources.

King-Hele has developed a meticulous editorial apparatus for the treatment of the chronologically presented sequence. Each letter is printed without the inclusion of any editorial distractions, according to his stated rules for transcription. Passages of some length are omitted in a few instances where these amount to an inserted "paper", such as details of Darwin's polygraphic machine and instructions on landscape drawing which have been removed from two 1779 letters. However, it is particularly pleasing to have facsimiles of the numerous sketches by means of which Darwin illuminated his observations, ideas and inventions. □

The editor's notes fill the spaces between the letters, and *fill* they sometimes do — 18 separate notes on a letter to Watt, and 100 lines of small-type notes on a 12-line letter to Joseph Banks. With a modest number to work on, King-Hele has been able to examine each letter under a powerful lens and reveal minutiae that might escape even the attentive reader's eye. Twenty years' research have brought him to an intimacy with his subject that could scarcely be equalled on the basis of surviving records, contemporary comment and astute interpretation. A dazzling breadth of erudition is woven into the fabric of his scores of enthusiastic and revealing notes. No stone, it seems, remains unturned and one admires his ability to meet the demands that the wealth of Darwin's ideas and activities has placed on him.

Precise information on the source, previous printings and text used for transcription, accompanies each letter, together with notes on the recipient, the date of the letter, explanations of obscurities and identification of people, places and things. A handbook on Darwin's life and times could be assembled from such detailed, well informed and bibliographically replete commentaries. Their biographical content alone is remarkable, for which the assistance of Hugh Torrens is acknowledged.

Fortunately for the reader, studiously careful guidance to all this scholarship is provided by means of name and subject indexes, a chronology, list of recipients and a gallery of 48 portraits, all the product of much thought and time. In fact the body of the text is so replete with editorial matter that it makes for slow reading if one progresses through letters and notes in turn.

As we have come to know from many previously published extracts, Darwin's letters are extremely engaging and entertaining reading, vividly reflecting his moods, interests, personal relations and doings. One really must read through runs of letters by themselves to get their flavour and then retrace one's steps for the additional pleasure of King-Hele's commentaries. Some of the letters are frankly dull, dealing with pedestrian business matters, and but for the overall scarcity of his correspondence might not have been included or accorded such reverential treatment. But most are packed with interest, varying from light-hearted banter to grave formality, and including verse, poignantly expressed sentiments, highly detailed technical descriptions, comments of immense foresight and wisdom, and reflections of his endless vigour, excitement with life and overwhelming good nature — in short, all that his biographers have taught us to expect. □

Gavin Bridson is Librarian and Archivist to the Linnean Society of London, and joint compiler of Natural History Manuscript Resources in the British Isles (Mansell, 1980).

The what, why and wherefore of cells

R.E. Stephens

Ultrastructure, Macromolecules, and Evolution. By Lawrence S. Dillon. Pp.708. ISBN 0-306-40528-8. (Plenum: 1981.) £43.79, \$69.50.

IN THIS second part of an intended trilogy, Lawrence Dillon interweaves discussion of evolutionary and cell structural relationships in order to consider genetic mechanisms working at the cellular level and to reconsider phylogeny from a cellular viewpoint. This is a formidable task because of the breadth and depth of information which must be assembled, assimilated and analysed. In terms of defining general principles, Dillon succeeds admirably; in terms of accurately documenting his background data, he fails in certain areas. Most of these failures are inconsequential to the ultimate conclusions, but they are disconcerting nevertheless.

The author has arranged his material into four areas of comparative cytology: cell motility, secretory organelles, energy-processing organelles, and karyo- and cytokinesis. He begins with a detailed account of the structural chemistry of membranes in general, giving an excellent historical perspective. Oddly, he follows this with four somewhat uneven chapters on cell motility before returning again to membrane-based systems. Motility is covered in chapters on microtubules and microfilaments, cytoplasmic movement, the flagellum and the basal apparatus. There are some serious problems in this area, to which I will return below.

Secretory systems are dealt with in two chapters — the endoplasmic reticulum and the Golgi apparatus are treated as a continuum, the "endomembrane system", while the lysosome, peroxisome and related structures are dealt with as the "vesicular system". This manner of presentation is quite informative in a comparative sense and it is logically and clearly arranged.

Three closely-correlated chapters are devoted to "energy-oriented organelles". The first discusses general aspects of cell respiration and includes a lucid consideration of the evolution of the tricarboxylic acid cycle; the remaining two cover mitochondria and chloroplasts, their comparative structure, replication and development in numerous organisms. Here, Dillon presents a well-balanced account of the two contending theories for the phylogenetic origin of these organelles: the endosymbiotic concept and the episome theory. He clearly favours the latter and presents additional evidence for it, based mainly upon some compelling evolutionary considerations.

Dillon ends this work with a comparison of both nuclear and cytoplasmic division

mechanisms in a multitude of organisms. Here, considering the abundance of information now available, I would have liked to see further examples, mention of odd exceptions and more detail, but the coverage is certainly adequate for the author's purposes. Tacked on to this final chapter are Dillon's two principal conclusions, concerning the inadequacy of simple mechanistic theory (i.e. the central dogma and self-assembly) to account for the development of cell structures and his well-stated contention that the phylogenetic stages of evolution, examined in terms of cellular evolution, allow no clear division between plants and animals, protozoans and algae, or even prokaryotes and eukaryotes. Certainly not everyone will agree with his views, but he raises many points that cannot be ignored.

The problems in the motility chapters, mentioned previously, range from simple inaccuracy (outer doublet B-subfibres contain 10 or 11 protofilaments, not 9; p.73 and p.143) and typographic atrocities (actinomysin for actinomycin, p.255; actinomyosin for actomyosin, p.407), to gross errors in reporting published conclusions (no analysis of doublet A and B tubulin components ever showed a specific number of residue differences, p.174) and unwarranted generalizations (in only one species has the ciliary membrane been shown to differ from the flagellar membrane, p.163). Also there are mistakes and misjudgements in citation priority, attributing the right facts to the wrong workers, and making overly broad generalizations from authors' more conservative conclusions. These, however, are no more than minor flaws which may easily be corrected in a second edition.

Perhaps more seriously, the dust jacket promises a new concept of flagellar movement. Dr Dillon argues against the widely-accepted sliding filament theory, concluding that it simply isn't so because microtubules are attached to the basal body. He begins his arguments by not mentioning Peter Satir's classic work on the subject, and by not considering that sliding is local and constrained and that the axoneme may twist in response to sliding. He concludes that the tubules themselves must change length, ignoring the work of D. P. Costello who argued all of this a decade ago and did so far more convincingly. The book would be much better off without this "new concept", particularly since it has nothing to do with Dillon's major premise.

Chemical differences in tubulins are central to Dillon's prime example for the workings of the cellular genetic mechanism (pp.174–176; p.558). With a simple self-assembly mechanism, how do the tubulins of the outer doublets and central pair sort

themselves out from each other and from cytoplasmic tubulin and, furthermore, when assembled into a 9 + 2 structure, how are dynein, nexin and spoke sites specified? The author contends that the mechanisms must be enzymatic. If the tubulins are really as different as Dillon implies, they should be able to sort themselves out easily (specific recognition sites) and also to specify binding sites when assembled (differences in surface lattice). In reality, however, the tubulins are not very different and some observed differences are probably due to post-translational modification, perhaps even after assembly. The truth of the mechanisms probably lies somewhere between these two viewpoints. This potentially important example sorely lacks authority and is likely to mislead the casual reader.

As an evolutionary exposition, *Ultrastructure, Macromolecules, and Evolution* is unique; nowhere is such a broad sweep presented so succinctly in a single source. The references, cited with full title and grouped by chapter at the end of the book, are quite valuable in their own right. This will be a useful reference book for those interested in comparative ultrastructure and certain phylogenetic aspects of biochemistry, and could also serve as a text for a graduate-level cell structure–function–evolution course, leading students easily back to the original source material. □

R.E. Stephens is Resident Investigator at the Marine Biological Laboratory, Woods Hole, Massachusetts, and Adjunct Professor of Physiology at the Boston University School of Medicine.

The healing way

N.R. Rowell

Tissue Repair and Regeneration. Handbook of Inflammation, Vol.3. Edited by L.E. Glynn. Pp.579. ISBN 0-444-80278-9. (Elsevier/North-Holland Biomedical: 1981.) \$136.50, Dfl.280.

REPAIR is essential to life, but what governs whether the body replaces damaged or dead tissue by tissue indistinguishable from the original, by scar or by a mixture of both, is a mystery. A distinguished international group of contributors here present a series of reviews concerning certain aspects of this problem. It is the third of five volumes dealing with inflammation, the first two of which dealt with chemical messengers and cell biology.

In the opening chapter Gabbiana and Runger-Brändle discuss the fibroblast, the fundamental cell in repair. From work in human beings and experimental animals, this cell has been shown to be remarkably adaptable and may alter its morphological, biochemical and functional features. In wound healing, fibroblasts, and possibly other cells, transform into contractile cells called myofibroblasts and these are probably concerned in wound contraction. Unfortunately, the authors cannot yet identify the factors controlling this process. Fibroblasts are also involved in the synthesis and metabolism of connective tissue, including collagen. Collagen has several functions as well as its primary role as a support. Duance and Bailey describe the five known types of collagen but admit that the relationship between structure and function is unknown. As many as 15 cell types may be involved in wound healing, including cells derived from the bone marrow, thymus, reticuloendothelial system and liver, as well as the cells of the tissues involved.

In regeneration, growth-regulating factors encourage the tissue to replace damaged areas with normal-functioning cells. Usually the formation of scar tissue is preceded by physical or other trauma, but this is not always the case; for example, in Dupuytren's contracture and scleroderma trauma is absent.

The book is nicely balanced between a consideration of basic processes and how these apply in the skin, central nervous system, lungs, heart and kidney. The editor himself has written an excellent chapter on the pathology of scar tissue, which links several of the authors' contributions. All authors agree that hyperplasia and atrophy, regeneration and repair are extremely complex phenomena, but in this book they have generally succeeded in making the subject intelligible to their readers.

Each chapter includes between 200 and 500 references which indicates the coverage of the subject. It also probably accounts for the fact that there are few references after 1978, although in one chapter additional notes have been added in 1980.

I can recommend this excellent book; it will be very useful for medical and non-medical scientists in many disciplines. Considering the complexity of the subject, the book is easy to read — presumably the result of sound editing — and there are few mistakes of any note and little overlap between contributions. Further volumes on the immunology of inflammation and the molecular basis for anti-inflammatory therapy will round off the series, which, when complete, should prove a most valuable addition to medical libraries in particular. □

N.R. Rowell is Reader in Dermatology at the University of Leeds, and Senior Consultant Physician in the Department of Dermatology, Leeds General Infirmary.

Modelling the growth of literature

B.C. Vickery

Scientific Information Systems and the Principle of Selectivity. By W. Goffman and K.S. Warren. Pp.191. ISBN 0-03-056081-0. (Praeger/Holt-Saunders: 1980.) £13.75, \$19.95.

IMPRESSED by the ever-growing rate of scientific publication, the authors of this book decided that the process of scientific communication needed to be studied as an "ecosystem". They develop a mathematical model for the growth of the literature in a defined subject, and then go on to a detailed quantitative analysis of various bodies of literature. The main experimental data consist of a bibliography of schistosomiasis, covering the period 1852–1962 and comprising over 10,000 references. There is little use made of the mathematical model in subsequent analysis, and there is no convincing demonstration of its validity.

Six kinds of analysis were made: distribution of articles and authors by language and by date; multiple publication by individual authors; multiple authorship of articles; most prolific journal titles and distribution of articles among journals; similar analyses for sub-areas of schistosomiasis research; and movement of researchers among sub-areas. The most interesting of the analyses is that based on co-authorship of articles. If authors A and B collaborate, also B and C, as well as C and D, then A,B,C,D can be said to form a network. Of the 6,500 authors in the schistosomiasis bibliography, 2,300 had no collaborators, but there was one big network that embraced 1,750 authors, as well as several hundred smaller networks.

This network of co-authorship is the only real evidence of scientific communication that is examined in the book. The citation of one author by another (implying awareness and probable reading of another's work) is not explored, still less any other forms of communication (attendance at conferences, for example).

The authors then conducted a qualitative evaluation of the schistosomiasis literature. Some 47 active workers in this subject were chosen with the aid of expert advisers. Each was asked to examine the bibliography and indicate "articles you deem of lasting importance". Over 3,000 of the 10,000 articles were selected at least once; about 1,600 at least twice; 471 at least five times; and 69 papers were selected 12 times or more. As one moved from the 3,000 to the 69, the likelihood of the authors being members of the large co-authorship network not surprisingly increased from 40 per cent to 70 per cent.

Goffman and Warren are convinced that "quality filtering" of scientific literature is needed in order to ease the task of retrieval, and offer proposals based on citation in review papers as an indicator of quality. I

find it doubtful whether any further mechanism — other than the review papers themselves — is needed by the scientific community.

The book presents a number of interesting ideas about the structure of scientific literature, and Goffman and Warren certainly extract as much as they can from the data used. But this information simply consists of the authors, title and source of each article, plus its evaluation by experts. There is no indication of what the authors have read, or who reads their contribution. In order to get a more rounded understanding of scientific communication, the book needs to be complemented by the reading of work such as that of Garvey, who has taken these other factors into consideration. □

Brian Vickery is Director of the School of Library Archive and Information Studies, University College London.

The complete angler

Bruce R. Moore

Circular Statistics in Biology. By Edward Batschelet. Pp.372. ISBN 0-12-081050-6. (Academic: 1981.) £28.80, \$69.50.

MANY scientists who rarely otherwise cite secondary sources, do so routinely when referring to statistical tests. Several reasons for this tradition are evident, not the least of which is that original statistics papers are meant to be read by statisticians.

This use of secondary sources is frequently quite harmless, since clear and accurate statistics texts are available at many levels. It has sometimes been less satisfactory, however, for that substantial minority of scientists whose data occur in vector form. The *t*-tests and ANOVAs of everyday statistics are seldom applicable to angular data. A number of special tests have, in fact, been developed, but they are not described in most statistics texts and are therefore missed by many potential users. In consequence, statistical inference based upon mere intuition has become the norm in certain areas.

Against this background, one finds isolated regions where the special tests have been thoroughly exploited. The source most often cited in these areas is an out-of-print monograph by Edward Batschelet (American Institute of Biological Sciences, 1965). Although extremely specialized, this work has become for many readers not only a highly trusted source but even an object of personal affection. Such persons were pleased to learn in recent years that, while approaching retirement, Professor

Batschelet was preparing a book-length expansion of the monograph.

Although he died unexpectedly a few months after retirement, Professor Batschelet's manuscript had been completed, and the book has now been published. Readers will find in it the same clarity of style and careful choice of example which characterized the earlier monograph. Essentially, Batschelet describes some 30 tests for vector clustering, group differences, correlations and goodness-of-fit. Most of these are briefly explained, then illustrated with one or more examples. Few derivations are given. The examples are taken almost entirely from biology, but readers from other disciplines should find them clear and easily translatable. Only slightly less clear are the supporting chapters describing various circular distributions and simple mathematical techniques.

There are a few mistakes: Example 6.5.1 requires $k = 80$; 7.4.1, $a_s = 2.25$; 7.7.1, $T^2 = 154$; p.183, $\Sigma \text{ sines} = -1.48$; and Table 9.3.2, a fresh beginning. The two-stage application of Mardia and Watson-

Wheeler tests in 7.8 is appropriate, but the sanctioned alternative combination of Mardia and Watson-Williams is not. The choice of median peak hour in 1.5.2 is not altogether proper, nor is the "careful" breaking of ties recommended in 9.3. But a few such lapses are almost inevitable; while unfortunate, they detract very little from the value of the book.

Circular Statistics should prove extremely helpful to almost anyone with vector data. It can be recommended also to potential writers of more general statistics texts. A worthwhile precedent is being set here by J. Zar's *Biostatistical Analysis*, the second edition of which, shortly to be published by Prentice Hall, includes two full chapters on vector-statistical methods. Batschelet offers an excellent starting point for a generation of other such chapters.

Altogether, we have received a most valuable legacy in this long-awaited and authoritative work. □

Bruce Moore is in the Department of Psychology at Dalhousie University, Halifax, Nova Scotia.

Life viewed in the particular

B. C. Goodwin

The Foundations of Biological Theory. By E. H. Mercer. Pp.242. ISBN 0-471-08797-1 (Wiley: 1981.) £29.20, \$22.55.

A PERENNIAL source of tension in science arises from the possibility of explaining phenomena in two distinct ways: either as a result of law, whereby they become intelligible in terms of principles at once simpler and more general than the phenomena themselves; or as a result of contingencies, of events that just happened to occur. A classic example of the latter was Buffon's suggestion that our planetary system arose from a chance event, a near-collision of the primitive Sun with another celestial body. Kant, on the other hand, suggested that orbiting planets arise by necessity (i.e. law) from gravitational instabilities in rotating masses of gas. Buffon the biologist chose a description in terms of particulars; Kant the philosopher sought a description in terms of universals. Current theory favours Kant, with far-reaching consequences, among them the expected existence of countless other planetary systems in the cosmos with conditions allowing life to evolve.

Biologists still favour description in terms of particulars, and this view dominates E. H. Mercer's clearly-written book on the conceptual foundations of biological theory. He assumes that all biological phenomena are instances of physical and chemical laws, but because of hereditary mechanisms, certain processes that arise by chance in organisms can be

perpetuated so that organisms become the sediments of historical contingencies which constrain in adaptively successful ways physical and chemical possibilities. For him "the historical course of events leading up to the present situation . . . provides biology with its theory". The task of theoretical biology is thus to tell the story of evolution, a point of view prevalent since Darwin. This has the important consequence that the biological realm is not rationally intelligible in the sense of revealing the operation of universal principles of organization and transformation. Life on other planets could look very different.

Mercer bases his analysis on the important distinction, established in Newtonian physics, between universal laws embodied in differential equations and the particular conditions which define a unique solution for a specific process. Although "boundary conditions" usually refer to particulars for fields, Mercer uses this term to include initial conditions, and he extends it to the more general concept of constraint, which simply imposes some limitation on the system. His treatment of biological organization is then presented in terms of a hierarchy of constraints which he describes as "organising or operational rules that define functional relationships between the units and regulate their interactions". The goal or purpose of biological organization is, he assumes, to ensure its survival. This is an extrinsic stability criterion. □

Many of the arguments are already familiar from the writing of Polanyi, Pattee, Riedl and others, to whom Mercer refers. Relevant ancillary concepts from such areas as information and control theory, thermodynamics and kinetics are clearly described, and these ideas are then used to describe organismic processes in terms of a hierarchy of constraints, with descriptive detail from biochemistry, genetics and development. A more analytical treatment in terms of non-equilibrium thermodynamics is less successful since Mercer claims that the near-equilibrium theory gives "a satisfactory account of the dynamic stability exhibited by adult organisms". However, such periodicities as circadian rhythms, neural pacemakers and peristaltic rhythms are excluded from this domain. Also, he incorrectly implies that a thermodynamic characterization has been given of these and other far-from-equilibrium phenomena; what has been obtained is useful insights into non-linear behaviour and symmetry-breaking by a number of investigators using stability and bifurcation theory, some of which Mercer describes.

Chance plays the conventional role of generating the variety on which natural selection operates, and the various sources from quantum indeterminacy to macroscopic uncertainty in bifurcations of "chaotic systems" and chromosome crossing-over are described. A final chapter outlining the origin and evolution of animate systems, continuous with physical evolution, completes the work. We are thus provided with a detailed account of organisms as constrained physico-chemical survival machines, both their hierarchical structure and their variety arising by chance.

Mercer's book has the virtues of clarity and extensive scholarship, only occasionally flawed by an incomplete grasp of subject matter. It provides an up-to-date account of neo-Darwinist thinking dominated by the "evolutionary paradigm", which defines the prevalent conceptual tradition in contemporary biology. However, there will continue to be those who, remembering Kant and seeing clear evidence of systematic regularity as well as variety in the biological realm, feel that Darwin's abandonment of explanation by law and his preoccupation with particulars (adaptation and inheritance), though historically understandable, was nevertheless misguided. For them, Mercer's approach will be unsatisfactory, and they will continue to seek a rational foundation for biology whereby organisms and their evolution become intelligible in terms of universal laws of organization and transformation, not simply in terms of chance events and survival. □

Brian Goodwin is Reader in Developmental Biology at the University of Sussex.